HARUZO HIDA

p-Adic Automorphic Forms on Shimura Varieties

Springer

Springer Monographs in Mathematics

Springer **M**onographs in **M**athematics

Springer
New York
Berlin
Heidelberg
Barcelona
Budapest
Hong Kong
London
Milan
Paris
Singapore
Tokyo

Haruzo Hida

p-Adic Automorphic Forms on Shimura Varieties

Springer

Haruzo Hida
Mathematics Department
UCLA
Box 951555
Los Angeles, California 90095

Mathematics Subject Classification (2000): 11F33, 11F41, 11F46, 11G18, 14G35

Library of Congress Cataloging-in-Publication Data
Hida, Haruzo.
p-adic automorphic forms on Shimura varieties / Haruzo Hida.
p. cm. — (Springer monographs in mathematics)
Includes bibliographical references and index.
ISBN 978-1-4419-1923-6 ISBN 978-1-4684-9390-0 (eBook)
DOI 10.1007/978-1-4684-9390-0
1. Shimura varieties. 2. Automorphic forms. 3. p-adic analysis. I. Title. II. Series.
QA242.5.H53 2004
515′.9—dc22 2003070362

ISBN 978-1-4419-1923-6 Printed on acid-free paper.

© 2004 Springer-Verlag New York, LLC
Softcover reprint of the hardcover 1st edition 2004

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer-Verlag New York, LLC, 175 Fifth Avenue, New York, NY 10010, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden. The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

9 8 7 6 5 4 3 2 1 SPIN 10950739

Springer-Verlag is a part of Springer Science+Business Media

springeronline.com

Preface

In the early years of the 1980s, while I was visiting the Institute for Advanced Study (IAS) at Princeton as a postdoctoral member, I got a fascinating view, studying congruence modulo a prime among elliptic modular forms, that an automorphic L-function of a given algebraic group G should have a canonical p-adic counterpart of *several* variables. I immediately decided to find out the reason behind this phenomenon and to develop the theory of ordinary p-adic automorphic forms, allocating 10 to 15 years from that point, putting off the intended arithmetic study of Shimura varieties via L-functions and Eisenstein series (for which I visited IAS). Although it took more than 15 years, we now know (at least conjecturally) the exact number of variables for a given G, and it has been shown that this is a universal phenomenon valid for holomorphic automorphic forms on Shimura varieties and also for more general (nonholomorphic) cohomological automorphic forms on automorphic manifolds (in a markedly different way).

When I was asked to give a series of lectures in the Automorphic Semester in the year 2000 at the Emile Borel Center (Centre Emile Borel) at the Poincaré Institute in Paris, I chose to give an exposition of the theory of p-adic (ordinary) families of such automorphic forms p-adic analytically depending on their weights, and this book is the outgrowth of the lectures given there.

In addition to an exposition of p-adic deformation theory of automorphic forms, I decided to give a comprehensive account of the construction of Shimura varieties (carrying a canonical family of abelian varieties) and Shimura's reciprocity laws characterizing them. Although this part is essentially expository, some new results of the author are included:

1. a (partially) new proof of some reciprocity laws of Shimura varieties;
2. irreducibility of the Igusa tower over unitary Shimura varieties;
3. a construction of invertible sheaves on Hilbert modular varieties that gives optimal p-integrality on Hilbert modular forms.

It is my hope that this book is useful to graduate students and researchers entering this area of research.

The author would like to thank the organizers of the Automorphic Semester at the Emile Borel Center for their invitation and also wishes to thank many people including his former and present graduate students who helped him to write correct English and mathematics. The author acknowledges partial support from the National Science Foundation (through the research grant: DMS 9988043 and DMS 0244401).

Los Angeles, California,
September 2003 *Haruzo Hida*

Contents

1 Introduction 1
1.1 Automorphic Forms on Classical Groups 5
1.2 p-Adic Interpolation of Automorphic Forms 8
1.3 p-Adic Automorphic L-functions 12
1.4 Galois Representations 13
1.5 Plan of the Book 13
1.6 Notation 15

2 Geometric Reciprocity Laws 17
2.1 Sketch of Classical Reciprocity Laws 18
2.1.1 Quadratic Reciprocity Law 18
2.1.2 Cyclotomic Version 19
2.1.3 Geometric Interpretation 20
2.1.4 Kronecker's Reciprocity Law 21
2.1.5 Reciprocity Law for Elliptic Curves 24
2.2 Cyclotomic Reciprocity Laws and Adeles 24
2.2.1 Cyclotomic Fields 24
2.2.2 Cyclotomic Reciprocity Laws 26
2.2.3 Adelic Reformulation 28
2.3 A Generalization of Galois Theory 31
2.3.1 Infinite Galois Extensions 31
2.3.2 Automorphism Group of a Field 35
2.4 Algebraic Curves over a Field 36
2.4.1 Algebraic Function Fields 36
2.4.2 Zariski Topology 43
2.4.3 Divisors 44
2.4.4 Differentials 45
2.4.5 Adele Rings of Algebraic Function Fields 50
2.5 Elliptic Curves over a Field 51
2.5.1 Dimension Formulas 51
2.5.2 Weierstrass Equations of Elliptic Curves 52

2.5.3 Moduli of Weierstrass Type 54
2.5.4 Group Structure on Elliptic Curves 55
2.5.5 Abel's Theorem 56
2.5.6 Torsion Points on Elliptic Curves 57
2.5.7 Classical Weierstrass Theory 60
2.6 Elliptic Modular Function Field 62

3 Modular Curves 67
3.1 Basics of Elliptic Curves over a Scheme 67
3.1.1 Definition of Elliptic Curves 68
3.1.2 Cartier Divisors 68
3.1.3 Picard Schemes 69
3.1.4 Invariant Differentials 70
3.1.5 Classification Functors 70
3.1.6 Cartier Duality 71
3.2 Moduli of Elliptic Curves and the Igusa Tower 72
3.2.1 Moduli of Level 1 over $\mathbb{Z}[\frac{1}{6}]$ 72
3.2.2 Moduli of $\mathcal{P}_{\Gamma_1(N)}$ 74
3.2.3 Action of $\mathbb{G}_m$ 75
3.2.4 Compactification 77
3.2.5 Moduli of $\Gamma(N)$-Level Structure 79
3.2.6 Hasse Invariant 80
3.2.7 Igusa Curves 81
3.2.8 Irreducibility of Igusa Curves 82
3.2.9 p-Adic Elliptic Modular Forms 85
3.3 p-Ordinary Elliptic Modular Forms 86
3.3.1 Axiomatic Treatment 86
3.3.2 Bounding the p-Ordinary Rank 89
3.3.3 p-Ordinary Projector 90
3.3.4 Families of p-Ordinary Modular Forms 90
3.4 Elliptic Λ-Adic Forms and p-Adic L-functions 92
3.4.1 Generality of Λ-Adic Forms 92
3.4.2 Some p-Adic L-Functions 94

4 Hilbert Modular Varieties 97
4.1 Hilbert–Blumenthal Moduli 98
4.1.1 Abelian Variety with Real Multiplication 98
4.1.2 Moduli Problems with Level Structure 102
4.1.3 Complex Analytic Hilbert Modular Forms 104
4.1.4 Toroidal Compactification 110
4.1.5 Tate Semi-Abelian Schemes with Real Multiplication . . 115
4.1.6 Hasse Invariant and Sheaves of Cusp Forms 117
4.1.7 p-Adic Hilbert Modular Forms of Level $\Gamma(N)$ 119
4.1.8 Moduli Problem of $\Gamma_1^1(\mathfrak{N})$-Type 123
4.1.9 p-Adic Modular Forms on $PGL(2)$ 125

4.1.10 Hecke Operators on Geometric Modular Forms 127
4.2 Hilbert Modular Shimura Varieties 131
4.2.1 Abelian Varieties up to Isogenies 133
4.2.2 Global Reciprocity Law 140
4.2.3 Local Reciprocity Law 154
4.2.4 Hilbert Modular Igusa Towers 158
4.2.5 Hecke Operators as Algebraic Correspondences 164
4.2.6 Modular Line Bundles 165
4.2.7 Sheaves over the Shimura Variety of $PGL(2)$ 175
4.2.8 Hecke Algebra of Finite Level 177
4.2.9 Effect on q-Expansion 178
4.2.10 Adelic q-Expansion 184
4.2.11 Nearly Ordinary Hecke Algebra with Central Character 189
4.2.12 p-Adic Universal Hecke Algebra 191
4.3 Rank of p-Ordinary Cohomology Groups 192
4.3.1 Archimedean Automorphic Forms 192
4.3.2 Jacquet–Langlands–Shimizu Correspondence 198
4.3.3 Integral Correspondence 202
4.3.4 Eichler–Shimura Isomorphisms 205
4.3.5 Constant Dimensionality 206
4.4 Appendix: Fundamental Groups 209
4.4.1 Categorical Galois Theory 209
4.4.2 Algebraic Fundamental Groups 216
4.4.3 Group-Theoretic Results 218

5 Generalized Eichler–Shimura Map 225
5.1 Semi-Simplicity of Hecke Algebras 225
5.1.1 Jacquet Modules 226
5.1.2 Double Coset Algebras 227
5.1.3 Rational Representations of G 230
5.1.4 Nearly p-Ordinary Representations 232
5.1.5 Semi-Simplicity of Interior Cohomology Groups 234
5.2 Explicit Symmetric Domains 236
5.2.1 Hermitian Forms over $\mathbb{C}$ 236
5.2.2 Symmetric Spaces of Unitary Groups 238
5.2.3 Invariant Measure 243
5.3 The Eichler–Shimura Map 244
5.3.1 Unitary Groups 245
5.3.2 Symplectic Groups 247
5.3.3 Hecke Equivariance 248

6 Moduli Schemes 251
6.1 Hilbert Schemes 251
6.1.1 Vector Bundles 252
6.1.2 Grassmannians 253
6.1.3 Flag Varieties 256
6.1.4 Flat Quotient Modules 258
6.1.5 Morphisms Between Schemes 262
6.1.6 Abelian Schemes 264
6.2 Quotients by $PGL(n)$ 269
6.2.1 Line Bundles on Projective Spaces 270
6.2.2 Automorphism Group of a Projective Space 270
6.2.3 Quotient of a Product of Projective Spaces 271
6.3 Mumford Moduli 276
6.3.1 Dual Abelian Scheme and Polarization 276
6.3.2 Moduli Problem 276
6.3.3 Abelian Scheme with Linear Rigidification 278
6.3.4 Embedding into the Hilbert Scheme 279
6.3.5 Conclusion 280
6.3.6 Smooth Toroidal Compactification 282
6.4 Siegel Modular Variety 290
6.4.1 Moduli Functors 291
6.4.2 Siegel Modular Reciprocity Law 293
6.4.3 Siegel Modular Igusa Tower 296

7 Shimura Varieties 303
7.1 PEL Moduli Varieties 304
7.1.1 Polarization, Endomorphism, and Lattice 304
7.1.2 Construction of the Moduli 309
7.1.3 Moduli Variety for Similitude Groups 314
7.1.4 Classification of G 316
7.1.5 Generic Fiber of $Sh_K^{(p)}$ 317
7.2 General Shimura Varieties 321
7.2.1 Axioms Defining Shimura Varieties 321
7.2.2 Reciprocity Law at Special Points 324
7.2.3 Shimura's Reciprocity Law 326

8 Ordinary p-Adic Automorphic Forms 329
8.1 True and False Automorphic Forms 329
8.1.1 An Axiomatic Igusa Tower 330
8.1.2 Rational Representation and Vector Bundles 331
8.1.3 Weight of Automorphic Forms and Representations 333
8.1.4 Density Theorems 335
8.1.5 p-Ordinary Automorphic Forms 339
8.1.6 Construction of the Projector e_{GL} 341
8.1.7 Axiomatic Control Result 344

8.2 Deformation Theory of Serre and Tate 345
8.2.1 A Theorem of Drinfeld 346
8.2.2 A Theorem of Serre–Tate 348
8.2.3 Deformation of an Ordinary Abelian Variety 349
8.2.4 Symplectic Case 350
8.2.5 Unitary Case 351
8.3 Vertical Control Theorem 355
8.3.1 Hecke Operators on Deformation Space 356
8.3.2 Statements and Proof 360
8.4 Irreducibility of Igusa Towers 363
8.4.1 Irreducibility and p-Decomposition Groups 364
8.4.2 Closed Immersion into the Siegel Modular Variety 364
8.4.3 Description of a p-Decomposition Group 367
8.4.4 Irreducibility Theorem in Cases A and C 369

References 375

Symbol Index 383

Statement Index 385

Subject Index 387

1

Introduction

The two major trends in number theory, automorphic and arithmetic, were initiated hand in hand in the mid nineteenth century by Kronecker and Kummer. One of Kronecker's legacies is the theory of arithmetic elliptic modular functions (and modular forms) having well-determined algebraic values (up to a specific transcendental factor: "period") at special points on modular curves. Shimura varieties (including modular curves) were principally studied by Shimura and Deligne in the later part of the last century. In particular, the theory gives a foundation of the rationality of automorphic L-values, because Shimura varieties often supply enough rationality for us to be able to pinpoint the transcendental factors of specific L-values (the automorphic periods).

On the other hand, the arithmetic side of number theory heavily influenced by Kummer's work is represented by Iwasawa's theory which interprets the algebraic part of the L-value as the size of a naturally given group (or module) defined by purely arithmetic means. The origin of this thought goes back to the class number formula of Dirichlet, Kummer, and Dedekind, which gives the size of the ideal class group of a number field as a special value of an abelian L-function (up to a power of the period $2\pi i$ of the exponential function). The main conjecture in Iwasawa's theory predicts precisely the same effect, replacing the class group by the so-called "Selmer group" and its "Iwasawa module" (attached to automorphic Galois representations) and abelian L-values by automorphic p-adic L-functions (see, for example, [MFG] Chapter 5).

Although these two trends seemed rather disparate until recently, the striking idea of Wiles (which has proven two of the most fundamental conjectures in number theory) has spanned a bridge between the two concepts. Indeed, it gives a way to find an automorphic form from a purely arithmetically given object, for example, a rational elliptic curve whose L-function is directly related to an elliptic modular Hecke L-function via the Shimura–Taniyama conjecture (now proven by Wiles and others; see [Wi2] and [BrCDT]).

This type of intricate relation between arithmetic and geometry can also be found in classical reciprocity laws. There are two aspects of the Artin reci-

procity law in class field theory. One is representation theoretic, for example, for the ring $\mathbb{A}^{(\infty)}$ of finite adeles,

$$\mathrm{Hom}_{cont}(\mathrm{Gal}(\mathbb{Q}^{ab}/\mathbb{Q}), \mathbb{C}^{\times}) \cong \mathrm{Hom}_{cont}((\mathbb{A}^{(\infty)})^{\times}/\mathbb{Q}_{+}^{\times}, \mathbb{C}^{\times})$$

via the identity of L-functions. Here $\mathbb{Q}^{ab}$ is the maximal abelian extension of $\mathbb{Q}$ (so, it is a composite of all Galois extensions of $\mathbb{Q}$ inside $\mathbb{C}$ with an abelian Galois group) and $\mathbb{Q}_{+}^{\times}$ is the group of positive rational numbers (see Theorem 2.12). Another geometric reciprocity can be formulated as

$$\mathrm{Gal}(\mathbb{Q}^{ab}/\mathbb{Q}) \cong GL_1(\mathbb{A}^{(\infty)})/\mathbb{Q}_{+}^{\times}.$$

Since this is a description of the field automorphism group $\mathrm{Gal}(\mathbb{Q}^{ab}/\mathbb{Q})$ by an adelic Lie group: $GL_1(\mathbb{A}^{(\infty)})$, we call it a geometric reciprocity law. They are equivalent by Pontryagin duality (which characterizes a locally compact abelian group by its characters), and the first is generalized (conjecturally) by Langlands in a non-abelian setting. Geometric reciprocity in a non-abelian setting would be via Tannakian duality. Tannakian duality characterizes a given (algebraic) group by the category of its rational representations with some additional data ([HMS] Chapter II and [Br]). A simple class of Shimura varieties (so-called, of PEL type) is given as a moduli variety of abelian schemes with a given endomorphism algebra (or equivalently, a moduli of motives of the first-degree cohomology of such abelian schemes). Such a motive H generates a Tannakian category whose dual group $MT_{H/\mathbb{Q}}$ (the so-called motivic Galois group of the motive) could be related in possibly different ways to the automorphism group of the Shimura variety. If the motive is sufficiently generic, the adele points of the group MT_H would give almost all automorphisms of the Shimura variety (cf. [D3] 1.7). Thus Shimura variety appears as a natural object of study if one wants general geometric reciprocity laws.

Iwasawa's theory is built upon the geometric reciprocity law. The cyclotomic field $\mathbb{Q}(\mu_{p^\infty})$ is the maximal p-ramified extension of $\mathbb{Q}$ fixed by $\widehat{\mathbb{Z}}^{(p)} \subset \mathbb{A}^{\times}/\mathbb{Q}^{\times}\mathbb{R}_{+}^{\times}$ removing the p-inertia multiplicative factor $\mathbb{Z}_p^{\times}$. We then try to study arithmetically constructed modules X (Iwasawa modules) from $\mathbb{Q}(\mu_{p^\infty}) \subset \mathbb{Q}^{ab}$; for example, Iwasawa himself studied the projective limit X (under the norm map) of the p-primary part of the class groups of number fields inside $\mathbb{Q}(\mu_{p^\infty})$. A principal idea is to regard X as a module over the Iwasawa algebra (which is a completed Hecke algebra relative to $\frac{GL_1(\mathbb{A}^{(\infty)})}{GL_1(\widehat{\mathbb{Z}}^{(p)})\mathbb{Q}_{+}^{\times}}$), and ring-theoretic techniques are used to determine X (see [ICF] Chapter 13).

General Tannakian duality produces from the category of all (pure) motives an enormous group, the so-called motivic Galois group (see [HMS] Chapters II through IV and [MTV]). However, for Iwasawa-theoretic applications, the full motivic Galois group appears to be hopelessly large to adapt. If one wants to create something like Iwasawa's theory in a non-abelian setting, we therefore need a scheme whose automorphism group has an identification with $G(\mathbb{A}^{(\infty)})/\overline{Z(\mathbb{Q})}$ for a reductive algebraic group G. If $G = GL(2)_{/\mathbb{Q}}$, the tower

$V_{/\mathbb{Q}^{ab}}$ of modular curves has $\mathrm{Aut}(V_{/\mathbb{Q}})$ identified with $GL_2(\mathbb{A}^{(\infty)})/Z(\mathbb{Q})$ as Shimura proved (see Theorem 2.43). The (geometric) decomposition group of (p) is given by $B(\mathbb{Q}_p) \times SL_2(\mathbb{A}^{(p\infty)})/\{\pm 1\}$ for a Borel subgroup B (see Corollary 4.22). We have been studying various arithmetically constructed modules over the Hecke algebra of $\frac{GL_2(\mathbb{A}^{(\infty)})}{GL_2(\widehat{\mathbb{Z}}^{(p)})U(\mathbb{Z}_p)\mathbb{Q}_+^\times}$, relative to the unipotent subgroup $U(\mathbb{Z}_p) \subset B(\mathbb{Z}_p)$ (removing the multiplicative inertia factor from the decomposition group). Often these are modules given by cohomology groups (either algebrogeometric or topological). Such study has yielded a p-adic deformation theory of p-ordinary automorphic forms (see Chapter 8), and it would be important to study the decomposition group at p of a given Shimura variety, which is close to the automorphism group of the mod p Shimura variety. We describe quite a few automorphism groups related to Shimura varieties of PEL type coming from symplectic and unitary groups (following Shimura's method and also using slightly more algebraic methods). In particular, the study of the p-decomposition subgroup provides us with a simple proof of the irreducibility of the Igusa tower over the mod p Shimura variety (see Section 8.4), which supplies us with one of the most useful tools in the p-adic study of automorphic forms: the q-expansion principle.

Besides determination of the automorphism groups of Shimura varieties, we present a concrete construction of p-integral models of Shimura varieties (following works of Shimura, Grothendieck, Mumford, Deligne, and Kottwitz) and the p-adic deformation theory of p-ordinary automorphic forms, in particular, the theory of analytic families of p-ordinary (cohomological) automorphic forms on reductive groups.

We are particularly interested in Shimura varieties that are canonically moduli varieties of abelian schemes with extra structures. Since abelian schemes are very specific projective subschemes of a projective space, we can study abelian schemes through classification theory of projective subschemes in a given projective space. The classification was initiated by W.-L. Chow and was completed by Grothendieck as the theory of Hilbert schemes. In the late 1950s (e.g., [59b] in [CPS] I), Shimura used the Chow coordinate associated with each projective variety in his determination of the field of moduli of a given family of abelian varieties (which was later incorporated as an essential ingredient in his construction of canonical models of automorphic varieties; see [AAF] 2.8 and Chapter II), and in [GIT] Mumford used Hilbert schemes to study the integral moduli of abelian schemes, which can be used to construct Shimura varieties. We give a down-to-earth exposition of the construction of the Hilbert schemes and the Mumford integral moduli in Chapter 6.

Assuming basic knowledge of abelian schemes (which can be found in [ABV]), toroidal compactifications of the moduli (in [DAV]), and class field theory, we then give an exposition of the construction of the Shimura varieties and their models over p-adic integer rings. The Shimura variety is characterized by its local and/or global *reciprocity* laws. The global reciprocity law gives a geometric expression of the automorphism group of the function field of the

Shimura variety and essentially determines the variety, because the Shimura variety is often a minimal model of the function field. The local reciprocity law determines via class field theory the move of special points, called *CM points*, under Galois action. The CM points are those points carrying abelian varieties with complex multiplication (also often called abelian varieties of *CM type*) and are Zariski dense in the variety, thereby providing rationality of the variety. An abelian variety has *complex multiplication* if its endomorphism algebra contains a semi-simple commutative algebra of degree over $\mathbb{Q}$ twice the dimension of the variety (see [ACM] 5.2 and [ABV] Section 19). The first principal purpose of this book is to supply the base of the construction of the (p-integral) Shimura variety in as elementary a manner as possible (including an exposition of the local and global reciprocity laws), because it is a pity that no other books (so far published) give a detailed scheme-theoretic construction of the variety now so important in every aspect of algebraic and geometric number theory. Shimura's books [AAF] and [EPE] give an elementary exposition on rationality of unitary and Siegel Shimura varieties and of automorphic forms and L-functions on them and will be indispensable for graduate students and researchers who want to enter this area of research.

After going through our exposition of Shimura varieties and p-adic automorphic forms on them, natural subjects of our exposition are

1. p-adic deformation theory of p-ordinary automorphic forms;
2. p-adic automorphic L-functions;
3. p-adic Galois representations;
4. the Iwasawa-theoretic significance of the p-adic L-functions.

In this book, we mainly touch upon the first topic with some examples of p-adic L-functions, leaving other topics to future treatments. Thus the second purpose of this book is to introduce integrality of automorphic forms on the varieties and hence to supply a foundation in order for researchers to pursue p-adic analytic study of automorphic forms and automorphic L-values.

Here is a brief history of p-adic modular and automorphic forms. The theory of p-adic elliptic modular forms was initiated by Serre in [Se3] taking the p-adic limit of q-expansions of elliptic modular forms. The geometric significance of Serre's definition was revealed by Deligne and Katz (e.g., see [K], [K1], and [K3]). Serre observed that the Eisenstein series of a given level Np has Fourier coefficients p-adically analytically depending on its weight k, and through this observation, a construction of p-adic zeta functions of totally real fields was given in [Se3], which was later brought to completion in [DR] through a more geometric formulation. A systematic study of elliptic cuspidal families of Hecke eigenforms p-adic analytically depending on weight was started by the author of this book in [H86a] and [H86b] and has been developed into a p-adic deformation theory of p-ordinary automorphic forms on Shimura varieties and classical groups (e.g., [H02], [H98], and [TiU]). More recently, Mazur, Gouvêa, and Coleman (e.g., [GM], [Co], and [Wa]) extended the theory to non-p-ordinary elliptic modular forms with positive slope. In par-

ticular, Mazur and Gouvêa conjectured, and Coleman has proven, that there always exists a p-adic analytic family of elliptic modular forms containing a given modular form with a positive slope, although only p-ordinary families seem to have a precise relation to classical objects (typical in Iwasawa's theory; see (3) in Section 1.2), because only for p-ordinary automorphic forms is p-adic analyticity of their coefficients described by Iwasawa's functions belonging to the Iwasawa algebra (and its finite flat extensions). So p-ordinary automorphic forms enjoy stronger p-adic analyticity than nonordinary forms. The deformation theory has found numerous profound applications by many different researchers. For example, it was used as an essential ingredient of a construction of two-dimensional Galois representations from Hilbert modular forms [Wi], of a proof of Iwasawa's conjecture in [Wi1], of the proof of Fermat's last theorem ([Wi2], [TaW], and see also [MFG] Chapter 3), of the proof of the Shimura–Taniyama conjecture ([Wi2], [BrCDT] and see also [GME] Chapter 5), and of the proof of the Artin conjecture for many nonsoluble icosahedral Galois representations (e.g., [BuT] and [BuDST]). More recently, the linear growth (relative to the degree of the base totally real fields) of the number of variables of the p-ordinary family of Hilbert modular forms was used by Skinner and Wiles ([SkW]) to prove modularity of residually reducible (but irreducible) two-dimensional p-adic Galois representations.

Since our second principal goal is to give an introductory exposition of p-adic automorphic forms, in the following sections of this introductory chapter, we give an overview of such automorphic forms (covering all four topics), putting off to the following chapter an introductory discussion on reciprocity laws (including a brief history).

1.1 Automorphic Forms on Classical Groups

Let $G_{/\mathbb{Q}}$ be a classical (reductive) group and $G_{1/\mathbb{Q}}$ be the derived group of G. Take a Borel subgroup $B_{/\mathbb{Q}}$ of G_1 and its torus $T_{/\mathbb{Q}} \subset G_1$. When G is split over $\mathbb{Q}$, we may embed G into $GL(n)_{/\mathbb{Q}}$ so that we may take B to be the group of upper triangular matrices in G_1 and T to be the group of diagonal matrices. We have $B = T \ltimes U$ for the unipotent radical U of B. On the quotient variety G_1/U (see Section 6.1.3), T acts by $gUt = gtU$, and hence T acts on the structure sheaf $\mathcal{O}_{G_1/U}$ by $t\phi(gU) = \phi(gtU)$. This action gives rise to an order on $X(T) = \mathrm{Hom}_{\text{alg. gp.}}(T_{/\overline{\mathbb{Q}}}, \mathbb{G}_m)$ so that the positive cone in $X(T)$ is made of κ such that the κ-eigenspace $L(\kappa)$ on the global sections of $\mathcal{O}_{G_1/U}$ is nontrivial. A character $\kappa \in X(T)$ is called the *weight* of T, and a weight is called *dominant* if it is in the positive cone. We then have a representation $L(\kappa; R) = L_G(\kappa; R)$ on $L(\kappa)$ given by $\phi(gU) \mapsto \phi(h^{-1}gU)$ for $h \in G_1(R)$, as long as T is split over a ring R. Often we can naturally extend the action of G_1 on $L(\kappa; R)$ to G (and we proceed hereafter in the introduction regarding $L_G(\kappa; R)$ as a G-module). When $G = GL(2)$, $T \cong \mathbb{G}_m$, $X(T) \cong \mathbb{Z}$ by $\kappa \leftrightarrow n$ if $\kappa(x) = x^n$, and $L(\kappa; R)$ is the κth symmetric tensor representation of $GL(2)$,

which can be realized on the space of homogeneous polynomials of degree n so that $\alpha \in GL(2)$ acts on a polynomial $P(X,Y)$ by $\alpha \cdot P(X,Y) = P((X,Y)^t\alpha^\iota)$ with $\alpha^\iota = \det(\alpha)\alpha^{-1}$.

Naively, an automorphic form on G is an $L(\kappa;\mathbb{C})$-valued analytic function on the symmetric space of the Lie group $G_1(\mathbb{R})$ with a prescribed move (or functional equation) under the action of a congruence subgroup Γ of $G(\mathbb{Q})$. A *congruence subgroup* of $G(\mathbb{Q})$ is a subgroup given by $S \cap G(\mathbb{Q})$ for an open compact subgroup S of $G(\mathbb{A}^{(\infty)})$ for the ring $\mathbb{A}^{(\infty)}$ of finite adeles (see Section 2.2.3 in the text and [MFG] Chapter 3 for adeles). A congruence subgroup is a discrete subgroup of $G_1(\mathbb{R})$ with finite covolume (i.e., $\int_{G_1(\mathbb{R})/\Gamma_1} dg < \infty$ for the Haar measure dg of $G_1(\mathbb{R})$ and $\Gamma_1 = \Gamma \cap G_1(\mathbb{Q})$). A congruence subgroup Γ is called *cocompact* if $G_1(\mathbb{R})/\Gamma_1$ is compact.

There are two ways of associating a weight with automorphic forms on G: one is to consider the cohomology group $H^d(\Gamma, L(\kappa;R))$ of an appropriate degree d for a given congruence subgroup $\Gamma \subset G(\mathbb{Q})$ and call harmonic automorphic forms spanning $H^d(\Gamma, L(\kappa;\mathbb{C}))$ automorphic forms of (topological) weight κ. This works well for any classical (or more general reductive) groups.

When the symmetric space of $G(\mathbb{R})$ is isomorphic to a (bounded) Hermitian domain $\mathcal{H}$ with origin $\mathbf{0}$, like (the restriction of scalars to $\mathbb{Q}$ of) F-forms of $GSp(2n)$ or $GU(m,n)$ over totally real fields F, we have another way to associate a weight with holomorphic automorphic forms. We have $\mathcal{H} \cong G_1(\mathbb{R})/C_{\mathbf{0}} = G(\mathbb{R})/C_{\mathbf{0}}Z(\mathbb{R})$ for the center Z of G and the stabilizer $C_{\mathbf{0}}$ in $G_1(\mathbb{R})$ of $\mathbf{0}$, which is a maximal compact subgroup of $G_1(\mathbb{R})$. In the case of $GL(2)_{/\mathbb{Q}} = GSp(2)_{/\mathbb{Q}}$, $C_{\mathbf{0}} = SO_2(\mathbb{R})$ and $\mathcal{H} = \mathfrak{H} := \{z \in \mathbb{C} | \operatorname{Im}(z) > 0\}$ with $G_1(\mathbb{R})/SO_2(\mathbb{R}) \cong \mathfrak{H}$ by $g \mapsto g(\sqrt{-1})$ (so, $\mathbf{0} = \sqrt{-1} \in \mathfrak{H}$). As is well known, $\mathfrak{H}$ is holomorphically equivalent to the open unit disk in $\mathbb{C}$ by $z \mapsto \frac{z-\sqrt{-1}}{z+\sqrt{-1}}$ (so it is equivalent to a bounded domain).

The group $C_{\mathbf{0}}$ can be regarded as a group of real points with respect to a twisted complex conjugation in the complexification C of $C_{\mathbf{0}}$. In the case of $GL(2)_{/\mathbb{Q}}$, $SO_2(\mathbb{R})$ can be considered as S^1 in $T(\mathbb{C}) = \mathbb{G}_m(\mathbb{C})$ by $\left(\begin{smallmatrix} * & * \\ c & d \end{smallmatrix}\right) \mapsto c\sqrt{-1} + d \in S^1$, and S^1 is the set of fixed points of the twisted "complex conjugation" $x \mapsto \overline{x}^{-1}$ in $T(\mathbb{C}) = \mathbb{C}^\times$. Generalizing this example, we see that the compact group $U(n)$ is the subgroup of $GL_n(\mathbb{C})$ fixed by the complex conjugation $x \mapsto {}^t\overline{x}^{-1}$. Any holomorphic representation $\rho : C \to GL(V(\rho))$ gives rise to a holomorphic complex vector bundle $\widetilde{V} = (G_1(\mathbb{R}) \times V)/C_{\mathbf{0}}$ by the action $(g,v) \mapsto (gu, u^{-1}v)$ for $u \in C_{\mathbf{0}}$. Since $\mathcal{H}$ is simply connected, we can split $\widetilde{V} \cong \mathcal{H} \times V$ as holomorphic vector bundles; so, we have a linear map $J_\rho(g,z) : V_z \to V_{g(z)}$ for each given $g \in G_1(\mathbb{R})$ which identifies the fibers V_z and $V_{g(z)}$ of $\widetilde{V}$. Thus we have a function $J_\rho : G_1(\mathbb{R}) \times \mathcal{H} \to GL(V)$ satisfying

1. (Cocycle Relation) $J_\rho(gh,z) = J_\rho(g,h(z))J_\rho(h,z)$ for $g,h \in G_1(\mathbb{R})$;
2. (Holomorphy) $J_\rho(g,z)$ is holomorphic in z and nonvanishing.

A function J_ρ as above is called an *automorphic factor* of G_1 (associated with the representation ρ). Automorphic factors of G are defined similarly.

When $G_1 = SL(2)_{/\mathbb{Q}}$, then $C_{\mathbf{0}} = SO_2(\mathbb{R}) \subset C = T(\mathbb{C})$ whose irreducible complex representation is given by

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} = \begin{pmatrix} * & * \\ c & d \end{pmatrix} \mapsto \rho \begin{pmatrix} * & * \\ c & d \end{pmatrix} = (ci+d)^k.$$

In this case, $J_\rho(g,z) = (cz+d)^k$. This goes as follows. Split $GL_2(\mathbb{R}) = PC_{\mathbf{0}}$ for P made of upper triangular matrices with 1 at the lower right corner. For $z = x+iy$, define $p_z = \left(\begin{smallmatrix} y & x \\ 0 & 1 \end{smallmatrix}\right)$. Then for $g \in SL_2(\mathbb{R})$, write $gp_z = p_{g(z)}u$ with $u \in C_{\mathbf{0}}$, and we have $\rho(u) = \rho(p_{g(z)}^{-1} g p_z) = (cz+d)^k$ by computation. Indeed, $J_\rho(g,z)$ sends (v, p_z) to $(uv, p_{g(z)}) \sim (v, gp_z) = (v, p_{g(z)}u)$.

One can view the complexification C as a real algebraic group; let T_C be a maximal real torus of C. With any character κ of T_C, we can associate a rational representation $\rho = \rho_\kappa^C : C \to GL(N)$ of C with representation space $L_C(\kappa; \mathbb{C})$ of dimension N. Let $V(\kappa) = L_C(\kappa; \mathbb{C})$. For a congruence discrete subgroup

$\Gamma \subset G(\mathbb{Q})$, a holomorphic automorphic form of (coherent) weight κ is a holomorphic function $f : \mathcal{H} \to V(\kappa)$ satisfying $f(\gamma(z)) = J_\rho(\gamma, z) f(z)$ for all $\gamma \in \Gamma$ (with some additional growth condition if $\Gamma\backslash\mathcal{H}$ is not compact). Again the space of holomorphic automorphic forms is trivial unless the weight κ is dominant (with respect to B chosen appropriately).

Often the complex manifold $\Gamma\backslash\mathcal{H}$ is canonically algebraizable, yielding an algebraic variety (or a scheme) X_Γ, called canonical models or *Shimura varieties*, defined over a valuation ring $\mathcal{W}$ in a number field with residual characteristic p. At the same time, we can algebraize the vector bundle $\widetilde{V}(\kappa)$ associated with $V(\kappa)$. Thus we can often have a coherent sheaf $\underline{\omega}^\kappa$ on X_Γ giving rise to $\widetilde{V}(\kappa)$ after extending the scalar to $\mathbb{C}$. The global sections of $H^0(X_\Gamma, \underline{\omega}^\kappa_{/R})$ for a $\mathcal{W}$-algebra R are called R-integral automorphic forms of weight κ. Note that T_C is isomorphic to T, because they are maximal tori in the same group G_1. Thus we identify T and T_C (with a compatible choice of Borel subgroups B and $B_C = B \cap C$). On X_Γ, we may regard the Γ-module $L_G(\kappa; R)$ as a locally constant sheaf associating, with an open subset $U \subset X_\Gamma$, sections over U of the covering space $\widetilde{L}(\kappa; R) = \Gamma\backslash(\mathcal{H} \times L_G(\kappa; R))$ over X_Γ. Here the quotient $\Gamma\backslash(\mathcal{H} \times L_G(\kappa; R))$ is taken through the diagonal action. Thus each dominant weight $\kappa \in X(T)$ gives two spaces of automorphic forms:

$$H^d(X_\Gamma, L_G(\kappa; R)), \quad G_\kappa(\Gamma; R) := H^0(X_\Gamma, \underline{\omega}^\kappa_{/R}).$$

There is (at least conjecturally) a correspondence $\kappa \mapsto \kappa^*$ such that

$$H^0(X_\Gamma, \underline{\omega}^\kappa) \hookrightarrow H^d(X_\Gamma, L_G(\kappa^*; \mathbb{C}))$$

by a generalized "Eichler–Shimura map" which is supposed to be equivariant under Hecke operators. If such equivariance holds, we say that the two modules, the source and the image, are equivalent as *Hecke modules*. In the example of $GL(2)_{/\mathbb{Q}}$, we have $\Gamma \subset SL_2(\mathbb{Z})$, $\kappa \in X(T) = X(\mathbb{G}_m) = \mathbb{Z}$, and $\kappa^* = \kappa - 2$ with

$$G_\kappa(\Gamma;\mathbb{C}) \hookrightarrow H^1(X_\Gamma, L_{GL(2)}(\kappa-2;\mathbb{C})),$$

and the original Eichler–Shimura map is given by sending f to the cohomology class $[f(z)(X-zY)^{\kappa-2}dz]$ of the holomorphic $L(\kappa-2;\mathbb{C})$-valued differential $f(z)(X-zY)^{\kappa-2}dz$ ([59c] and [62a] in [CPS], [IAT] Chapter 8, and see also [LFE] Chapter 6). This is valid if $\kappa \geq 2$.

1.2 p-Adic Interpolation of Automorphic Forms

We now interpolate these two families of cohomology groups

$$\{H^0(X_\Gamma, \underline{\omega}^\kappa_{/R})\}_\kappa \quad \text{and} \quad \{H^d(X_\Gamma, L_G(\kappa;R))\}_\kappa$$

when the weights κ vary continuously in the p-adic analytic group $\mathcal{D} = \mathrm{Hom}_{\text{top-gp}}(T(\mathbb{Z}_p), \overline{\mathbb{Q}}_p^\times)$ (fixing an algebraic closure $\overline{\mathbb{Q}}_p$ of $\mathbb{Q}_p$). Here to describe our idea of how to interpolate automorphic forms, we write W for the p-adic completion of $\mathcal{W}$.

On these two families of spaces, there is a natural action of Hecke operators; so, we also want this interpolation to take into account the Hecke operators. To carry out the p-adic interpolation, we need a good p-integral structure on the group G; so, we hereafter assume that G is defined over $\mathbb{Z}_{(p)} = \mathbb{Q} \cap \mathbb{Z}_p$ as an affine group scheme and that G_1 is isomorphic over $\mathbb{Z}_p$ to the restriction of the scalar to $\mathbb{Z}_p$ of a classical Chevalley group over the p-adic integer ring O_p of an unramified finite semi-simple extension of $\mathbb{Q}_p$; so, the Borel subgroup B and its maximal torus T is well defined over $\mathbb{Z}_{(p)}$, and T is split over O_p. Thus, if $O_p = \mathbb{Z}_p$, $T(\mathbb{Z}_p) \cong (\mathbb{Z}_p^\times)^r$ for the rank r of G_1 over $\mathbb{Q}_p$, and $\mathcal{D}$ is (isomorphic to) a disjoint union of finitely many copies of the p-adic (open) unit disk. In general, $T(\mathbb{Z}_p)$ is a product of a torsion-free multiplicative group Γ_T isomorphic to the additive group $\mathbb{Z}_p^n$ and a finite group. Then the *Iwasawa algebra* $W[[T(\mathbb{Z}_p)]]$ (which is the projective limit of the group algebras $\{W[T(\mathbb{Z}/p^m\mathbb{Z})]\}_m$) is a free module of finite rank over $W[[\Gamma_T]]$, which is in turn (noncanonically) isomorphic to the power series ring $W[[T_1,\dots,T_n]]$ of n variables. What we now do is:

(1) (Universality) Construct a (big) compact module V over $W[[T(\mathbb{Z}_p)]]$ such that the κ-eigenspace $V[\kappa]$ contains canonically, as $W[[T(\mathbb{Z}_p)]]$-submodules, $H^d(X_\Gamma, L(\kappa;W))$ in the topological case, and $H^0(X_{\Gamma/W}, \underline{\omega}^\kappa)$ in the coherent case, respectively;
(2) (Hecke operators) Establish a natural action of Hecke operators on V, and show the inclusion in (1) is Hecke equivariant;
(3) (Precise control) Find an appropriate $W[[T(\mathbb{Z}_p)]]$-submodule $X \subset V$ of cofinite type ($\Leftrightarrow$ W-dual is of finite type) such that X is stable under Hecke operators and $X[\kappa]$ (if $\kappa \gg 0$) is canonically isomorphic to a well-described subspace of automorphic forms of weight κ as Hecke modules.

Item (3) is often called a *vertical control theorem* (abbreviated as VCT) of the subspace X. Contrary to the case of complex-valued automorphic forms, p-adic automorphic forms, as a whole, regarded as a representation space of $G(\mathbb{A}^{(p\infty)})$, have little discrete spectrum, and they tend to involve big continuous spectrums. A vertical control theorem cuts out from the entire space of p-adic automorphic forms a continuous spectrum well behaving under the toric action. Examples of VCT are given as Theorem 3.7 for elliptic modular forms, Theorem 3.8 for a p-adic family of elliptic modular forms, Theorem 4.9 for Hilbert modular forms, Corollary 4.32 for Hilbert modular Hecke algebras and Theorem 8.13 for automorphic forms on unitary groups. A more general result on VCT can be found in [H02].

We mainly deal with the coherent case where G admits Shimura varieties that are given as moduli of abelian varieties with level structure (the so-called PEL structure: a polarization and points of finite order, a specified endomorphism algebra, and a specific lattice giving integral homology of abelian varieties). We give an exposition of the construction of the p-integral models of the variety in Chapters 6 and 7. However, we need to use some results obtained from the topological consideration; so, some sections in Chapter 5 are also devoted to describing the situation in topological cases.

In the coherent case, we define V to be the space of p-adic automorphic forms that classify abelian schemes with "ordinary" level p^∞ structure in addition to a PEL structure outside p. In other words, V is the space of formal functions on a formal (pro-)scheme, called the Igusa tower, which is the formal modulus classifying abelian schemes over p-adically complete W-algebras with "ordinary" level p^∞-structure. We prove a vertical control theorem for the space $X = V^{ord}$ of nearly p-ordinary automorphic forms and prove that its W-dual $\mathrm{Hom}_W(V^{ord}, W)$ is $W[[T(\mathbb{Z}_p)]]$-projective of finite type.

Actually, we can give for any classical group a good definition of nearly p-ordinary cusp forms; that is, a cusp form is called nearly p-ordinary if the Newton polygon of its Hecke polynomial at p is equal to the hypothetical Hodge polygon mechanically constructed from the weight κ (of the motive attached to the cusp form). We prove that the Newton polygon is always on or above the Hodge polygon (without recourse to hypothetical motives); so, a nearly p-ordinary form has the lowest Newton polygon identical to the Hodge polygon (see Chapter 5). For a Hecke eigenform of high weight (i.e., the weight is inside the positive cone of dominant weights far from its boundaries), whether it is nearly p-ordinary for almost all primes p is a difficult question to answer (cf. [Gu]). If the eigenform has low weight (conjectured to be associated with Artin Galois representations or abelian varieties), the density 1 property of ordinary primes is expected, and it is known for elliptic modular forms of weight 2 (see [Se4] Section 7) and is obvious for Artin representations. Thus lifting such a low weight p-ordinary form to a family of nearly p-ordinary automorphic forms is at present the only theoretical way to produce high-weight nearly p-ordinary forms.

One would expect that $\widehat{V}^{ord} := \mathrm{Hom}_W(V^{ord}, W)$ is $W[[T(\mathbb{Z}_p)]]$-projective of finite rank if G admits Shimura varieties. Contrary to this, when we deal with a group such as $GL(n)$ $(n > 2)$ without having Shimura variety, the module $\widehat{V}^{ord}$ is of finite type over $W[[T(\mathbb{Z}_p)]]$, but we can prove that it is of torsion (although there is a conjectural formula of its Krull dimension; see [H95] and [H98]). Natural questions are:

(Q1) *When can one expect that the space $\widehat{V}^{ord}$ is $W[[T(\mathbb{Z}_p)]]$-projective? What is the (expected) minimal value of κ at which the vertical control holds? What happens if one specializes at very low weight? If $\widehat{V}^{ord}$ is torsion, what is the Krull dimension of the $W[[T(\mathbb{Z}_p)]]$-module $\widehat{V}^{ord}$? What is its characteristic power series if the codimension of $\widehat{V}^{ord}$ equals 1 in $Spec(W[[T(\mathbb{Z}_p)]])$?*

It turns out that all these questions are quite arithmetic, as we show in the principal text. In the elliptic modular case, the lowest weight where VCT holds is 2. However, as Buzzard and Taylor studied in [BuT], there is a good criterion via Galois representations to guarantee the limit at weight 1 to be a true modular form (not just p-adic), which played an important role in the proof of the Artin conjecture for some icosahedral cases by the school of R. Taylor ([SBT], [Ta1], and [Ta2]).

Write $X[p^n]$ for the kernel of multiplication by p^n of an abelian scheme X (as a locally free group scheme over the base), and define its Barsotti–Tate group (p-divisible group) by $X[p^\infty] = \varinjlim_n X[p^n]$. Our description of automorphic forms is geometric, whose origin goes back to some work of Katz and Deligne–Ribet (cf. [K3] and [DR]). In the example of $G = GL(2)_{/\mathbb{Q}}$, we take an arbitrary p-adically complete W-algebra $R = \varprojlim_n R/p^nR$, and consider a test object $(E, \phi_p, \phi_N)_{/R}$ made of an elliptic curve E, a level p^∞-structure $\phi_p : \mu_{p^\infty} \hookrightarrow E[p^\infty]$ (i.e., a closed immersion of ind-group schemes) and a level N-structure ϕ_N, like a point of order N (an inclusion of $\mathbb{Z}/N\mathbb{Z}$ into the set of N-torsion elements $E[N]$ in E), all these data being defined over R. A *p-adic elliptic modular form* f is a functorial rule associating an element of R to a test object $(E, \phi_p, \phi_N)_{/R}$. Thus we have $f(E, \phi_p, \phi_N) \in R$, and for each p-adically continuous W-algebra homomorphism $R \xrightarrow{\rho} R'$, we have

$$f((E, \phi_p, \phi_N)_{/R,\rho} \times R') = \rho(f(E, \phi_p, \phi_N)).$$

Though f is a functorial morphism (as we said a "rule"), hereafter abusing the language, we pretend as if f were a function of test objects (obeying the specific rules we have described). A modular form which is an eigenform of $T(p)$ (or $U(p)$) is *p-ordinary* if its eigenvalue at p is a p-adic unit. In general, p-ordinary modular forms are linear combinations of such (we give a more thorough definition in the text; see Section 3.3). The evaluation of f at the Tate curve $Tate(q)_{/\mathbb{Z}((q))}$ (see Section 3.2.4) at the cusp ∞ yields the q-expansion:

$$f(q) = f(Tate(q)_\infty, \phi_p^{can}, \phi_N^{can}) = \sum_{n=0}^{\infty} a(n,f)q^n.$$

We can deduce from the irreducibility of the Igusa tower that V is the p-adic completion of

$$\mathcal{W}[[q]] \bigcap \left(\sum_{k=0}^{\infty} \sum_{\alpha} G_k(\Gamma_1(Np^\alpha); \mathbb{C}) \right).$$

We may identify the Fourier expansion of $f \in G_k(\Gamma_1(N); \mathbb{C})$ with the q-expansion $f(q) \in \mathbb{C}[[q]]$, writing $q = \exp(2\pi iz)$.

In the topological case of $G = GL(2)_{/\mathbb{Q}}$, V is given by the cohomology group $H^1(\Gamma', \mathcal{C}(\widehat{\Gamma}'/U(\mathbb{Z}_p), W))$, where $\mathcal{C}(\widehat{\Gamma}'/U(\mathbb{Z}_p), W)$ is the space of continuous functions on the p-adic analytic space $\widehat{\Gamma}'/U(\mathbb{Z}_p)$, $\Gamma' = \Gamma_1(N) \cap \Gamma_0(p)$ and $\widehat{\Gamma}'$ is the closure of Γ' in $GL_2(\mathbb{Z}_p)$. Then $\widehat{\Gamma}' = \widehat{\Gamma_0(p)}$, and we may regard $P \in L(\kappa; W)$ as a polynomial function on $\widehat{\Gamma}_0(p)/U(\mathbb{Z}_p) \subset (SL(2)/U)(\mathbb{Z}_p)$, and this inclusion, $L(\kappa; W) \hookrightarrow \mathcal{C}(\widehat{\Gamma}_0(p)/U(\mathbb{Z}_p), W)$, induces the map $H^1(\Gamma', L(\kappa; W)) \hookrightarrow V$.

For any $W[[T(\mathbb{Z}_p)]]$-module X satisfying VCT, the eigenvalue $\lambda(t)$ of each Hecke operator t acting on X is algebraic over $W[[T(\mathbb{Z}_p)]]$. Indeed, the Hecke algebra $\mathbf{h}$ in $\mathrm{End}_{W[[T(\mathbb{Z}_p)]]}(X)$ generated by (appropriate) Hecke operators is an algebra over $W[[T(\mathbb{Z}_p)]]$ of finite (generic) rank (or even of torsion). Take an irreducible component $Spf(\mathbb{I})$ of the formal spectrum $Spf(\mathbf{h})$ of $\mathbf{h}$. The operator t projected to $\mathbb{I}$, written as $\lambda(t)$ (i.e., $\lambda : \mathbf{h} \twoheadrightarrow \mathbb{I}$ is the projection), can be considered to be a formal function on the formal scheme $Spf(\mathbb{I})$. The p-adic space $Spf(\mathbb{I})(W)$ of W-points of $Spf(\mathbb{I})$ is given by the set of continuous W-algebra homomorphisms of $\mathbb{I}$ into W. In particular, if $P \in Spf(\mathbb{I})(W)$ satisfies $P|_{W[[T(\mathbb{Z}_p)]]} = \kappa$ for $\kappa \gg 0$, $\lambda(t)(P) = P(\lambda(t))$ is the eigenvalue of t occurring in either $H^0(X_\Gamma, \underline{\omega}^\kappa)$ or $H^d(X_\Gamma, L(\kappa; W))$.

When $G = GL(2)_{/\mathbb{Q}}$, $T(\mathbb{Z}_p) = \mathbb{Z}_p^\times = u^{\mathbb{Z}_p} \times \Delta$ for a finite group Δ. Thus $W[[T(\mathbb{Z}_p)]] = \Lambda[\Delta]$ for $\Lambda = W[[u^{\mathbb{Z}_p}]] \cong W[[U]]$ (the classical Iwasawa algebra and a formal power series ring) via $u^s \mapsto (1+U)^s = \sum_{n=0}^{\infty} \binom{s}{n} U^n$. Note that

$$\kappa((1+U)^s) = \kappa(u^s) = u^{\kappa s} = (1+U)^s|_{U=u^\kappa - 1}.$$

The algebra homomorphism $\kappa : \Lambda \to W$ is the "evaluation" at $U = u^\kappa - 1$! Thus if $\mathbb{I} = \Lambda$, $Spf(\Lambda)(\overline{\mathbb{Q}}_p)$ is the open unit disk D° in $\overline{\mathbb{Q}}_p$ by $\kappa \mapsto \kappa(u) - 1$, and $\lambda(T(n))(\kappa) = \lambda(T(n))(u^\kappa - 1)$ (viewing $\lambda(T(n))$ as a power series) gives a p-adic analytic interpolation of Hecke eigenvalues over D°. This type of p-adic analytic function $s \mapsto \Phi(u^s - 1, \ldots, u^s - 1)$ for a power series $\Phi(T_1, \ldots, T_n) \in W[[T_1, \ldots, T_n]]$ is called an *Iwasawa function* (which gives a special subclass of p-adic analytic functions on the product of copies of the open unit disk D°). In general, we get the p-adic interpolation of Hecke operators parameterized by $Spf(\mathbb{I})(W)$.

1.3 p-Adic Automorphic L-functions

Fix embeddings $i_\infty : \overline{\mathbb{Q}} \to \mathbb{C}$ and $i_p : \overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_p$. Since the specialization $\lambda(T(n))(u^\kappa - 1)$ (in the simplest case of $GL(2)$) is the Hecke eigenvalue occurring in the space of cusp forms, it can be considered as a complex number uniquely. Thus we can think of automorphic L-functions $L(s, \lambda(\kappa))$ made from such eigenvalues; for example, the modular Hecke L-function

$$L(s, \lambda(\kappa)) = \sum_{n=1}^{\infty} i(\lambda(T(n))(u^\kappa - 1))n^{-s},$$

writing $i = i_\infty \circ i_p^{-1}$.

In general, for sufficiently positive weight κ of T, we have Hecke eigenvalues $\lambda(t)(\kappa)$ occurring on the coherent cohomology with coefficients in $\underline{\omega}^\kappa$ on the Shimura variety. We thus have a family $\{\lambda(\kappa)\}_\kappa$ of systems of eigenvalues (or equivalently, a family of holomorphic automorphic representations). We give ourselves a rational representation r of the Langlands dual ${}^L G$ of G and consider the family of automorphic L-functions $\{L(s, \lambda(\kappa), r)\}_\kappa$. Suppose that $L(m, \lambda(\kappa), r)$ for (sufficiently many) integers m (depending on κ) has rationality up to a transcendental factor or a period $\Omega(m, \kappa, r)$. Then natural questions we would ask are:

(Q2) *Is it possible to interpolate p-adically the value* $\frac{L(m,\lambda(\kappa),r)}{\Omega(m,\kappa,r)}$ *? Is it possible to find* $L_{\lambda,r}(U) \in \mathbb{I}[[U]]$ *such that the p-adically analytic function*

$$\mathbb{Z}_p \times Spf(\mathbb{I})(W) \ni (s, P) \mapsto L_p(s, P, r) = P(L_{\lambda,r}(u^s - 1))$$

satisfies the identity: $L_p(m, \kappa, r) = \frac{L(m,\lambda(\kappa),r)}{\Omega(m,\kappa,r)}$ *(up to a standard modification factor) for points* $(m, \kappa) \in \mathbb{Z} \times Spf(\mathbb{I})(W)$ *with* $\kappa|_{T(\mathbb{Z}_p)} \in X(T)$ *densely populated in* $\mathbb{Z}_p \times Spf(\mathbb{I})(W)$*?*

This problem of course involves a subtle integrality question of how to normalize the factor $\{\Omega(m, \kappa, r)\}_{\kappa,m}$ in the aggregate (varying $\kappa \in X(T)$ and integers m) to obtain an "optimal" integrality; so, it is more involved than proving rationality (see Section 3.4 for some examples and [SGL] for a general theory). Once we are successful in constructing canonical p-adic L-functions, we could ask more specifically,

(Q3) *When is the p-adic L-function entire? Where could it have singularities? If there is a singularity, what is the order of the zero or the pole? If we have a pole, what is the residue?*

See [SGL] for particular G and r for which we have a partial answer to these questions and for some conjectural discussion on these questions.

1.4 Galois Representations

Once an irreducible component $\mathbb{I}$ of the Hecke algebra is given, one would expect to have a Galois representation $\rho_\lambda : \mathrm{Gal}(\overline{\mathbb{Q}}/F) \to {}^L G(\mathbb{I})$ such that the Hecke polynomial at $\ell \neq p$ gives rise to the characteristic polynomial of the Frobenius element. Here abusing our language, ${}^L G$ is the p-integral Langlands dual defined in an appropriate manner (see [DGH]). We could heuristically (or conjecturally) answer the question (Q3) that the associated p-adic L-function $L_p(s, P, r)$ has singularity at P if the Galois representation $r \circ \rho_\lambda$ specialized at P; that is, $(r \circ \rho_\lambda) \mod P$ contains the trivial representation (a p-adic Artin conjecture; see [SGL]). We then further ask

(Q4) *For a pair of representations (ρ_λ, r) as above, is there any good way to associate a Selmer group $\mathrm{Sel}(r \circ \rho_\lambda)$ so that the characteristic element in $\mathbb{I}$ of the Selmer group should be equal to the p-adic L-function or its numerator?*

See [MFG] Chapter 5 for a general description of Selmer groups. If this is affirmative, then this gives a description of the zero-set of the p-adic L-function in arithmetic terms. Related to this, VCT is often useful to identify the nearly p-ordinary Hecke algebra with the nearly p-ordinary universal deformation ring if at one weight the deformation ring with the fixed weight is identified with the Hecke algebra of the specific weight (see [Wi2] Theorem 3.3, [MFG] 5.3.5, and [HM] Section 4.3 for this type of argument in the case of $GL(2)/\mathbb{Q}$). The argument proving VCT often yields another type of control theorem: the so-called horizontal control theorem (HCT; e.g., [GME] 3.1.2 and 3.2.5), giving a precise description of the behavior of a Hecke algebra if one adds primes outside p to the level of the Hecke algebra. This horizontal control (HCT) is used in the case of $GL(2)$ to construct the Taylor–Wiles systems which was invented by Wiles and Taylor in the celebrated papers [Wi2] and [TaW]. With the help of the system, Wiles obtained the identification of the Hecke algebra of a specific weight with the Galois deformation ring (an essential ingredient of Wiles' proof of the Shimura–Taniyama conjecture and Fermat's last theorem). See [Wi2], [MFG] Chapter 3, and [GME] Chapter 3 for these topics.

1.5 Plan of the Book

We answer some of the above questions in this book in some specific cases in a concrete way and in some other cases conjecturally. Since this book is an outgrowth of a series of lectures given at the Emile Borel Center in its Automorphic Semester of the year 2000, it would be helpful to include here the original plan of the lectures to give a brief outline of the structure of the book. Here is the plan.

1. In the first two chapters (Chapters 2 and 3), we recall the classical reciprocity law and its generalization in the theory of the elliptic modular

forms. Chapter 2 is from the notes of a graduate course I gave at UCLA in the academic year 1999–2000, and from Chapter 3 on, the book is an expanded version of the lectures at the Emile Borel Center (see [H03a] for the original lecture notes). We also give a brief proof of the VCT in the elliptic modular case as a prototype of the general theory.

2. Chapter 4 is devoted to another example: Hilbert modular forms. There we prove diverse vertical control theorems as well as the freeness of the universal nearly ordinary Hecke algebra over the Iwasawa algebra of $T(\mathbb{Z}_p)$. We also give fairly detailed proofs of global and local geometric reciprocity laws for Hilbert modular functions, which are slightly more algebraic than Shimura's treatment (although the basic ideas behind the proof are the same).
3. In Chapter 5, we define near-ordinarity of cohomological automorphic forms by comparing the p-adic Newton polygon and the Hodge polygon of Hecke eigenforms. Furthermore we prove semi-simplicity of the Hecke operator action on topological nearly ordinary cohomology groups (for general reductive groups).
4. In Chapter 6, we recall the basic techniques (due to Grothendieck and Mumford) of constructing algebrogeometric moduli schemes. We state and prove, at the end, Shimura's global reciprocity law for Siegel modular functions.
5. In Chapter 7, we first construct the Shimura varieties via the integral moduli theory of abelian schemes. Thereafter we recall briefly the global reciprocity law for general Shimura varieties not necessarily of PEL type (without any detailed proof).
6. In Chapter 8, we prove the VCT for unitary and symplectic groups. A key ingredient is the deformation theory (due to Serre-Tate) of p-ordinary abelian schemes. At the end, combining the global reciprocity law and the deformation theory, we prove the irreducibility of the Igusa tower for unitary and symplectic groups. This supplies us with the q-expansion principle for Shimura varieties of symplectic and unitary groups. We should point out that the computation of cuspidal monodromy which successfully proved the irreducibility of the Igusa tower in the Siegel modular case ([DAV] V.7) does not yield the irreducibility for groups of the form $\mathrm{Res}_{F/\mathbb{Q}}G_0$ for unitary and symplectic groups G_0 over a totally real field F bigger than $\mathbb{Q}$ by a reason explained in [DT] Section 12 (see also a remark after Theorem 4.21). This point is subtly related to the Leopoldt conjecture for F and p.

The author hopes to write another book (expanding Chapter 4) to cover the theory of automorphic p-adic L-functions, horizontal control theorems, and the construction of Galois representations in the Hilbert modular case, which have been very useful in proving some main conjectures in the (generalized) Iwasawa theory and in identifying Hecke algebras treated in this book with

Galois deformation rings (with a specific universal property; see [F1], [SkW], and [MFG]).

Some of the papers and preprints of mine related to these subjects can be downloaded from my web site: www.math.ucla.edu/~hida.

1.6 Notation

In this book, we continue to use the notation utilized in this introduction. Here we describe some of the notation which is used without much explanation. The symbol $\mathbb{Z}_p$ denotes the p-adic integer ring inside $\mathbb{Q}_p$, and the symbol $\mathbb{Z}_{(p)}$ is used to indicate the valuation ring $\mathbb{Z}_p \cap \mathbb{Q}$. A subfield E of $\overline{\mathbb{Q}}$ is called a number field (often assuming $\dim_{\mathbb{Q}} E < \infty$ tacitly). For a number field E, O_E denotes the integer ring of E, $O_{E,p} = O_E \otimes_{\mathbb{Z}} \mathbb{Z}_p \subset E_p = E \otimes_{\mathbb{Q}} \mathbb{Q}_p$ and $O_{E,(p)} = O_E \otimes_{\mathbb{Z}} \mathbb{Z}_{(p)} \subset E$. The symbol $\mathcal{W}$ is exclusively used to indicate a valuation ring inside $\overline{\mathbb{Q}}$ with residual characteristic p, and often $\mathcal{W}$ is supposed to be unramified over $\mathbb{Z}_{(p)}$. The p-adic completion $\varprojlim_n \mathcal{W}/p^n\mathcal{W}$ is denoted by W, and we write $W_m = W/p^m W = \mathcal{W}/p^m\mathcal{W}$.

The symbol $\mathbb{A}$ is the adele ring of $\mathbb{Q}$, and for a vector space of a number field E, we write $V_{\mathbb{A}} = V(\mathbb{A})$ for $V \otimes_{\mathbb{Q}} \mathbb{A}$. For a subset Σ of the set of rational places, we set $\mathbb{A}^{(\Sigma)} = \{x \in \mathbb{A} | x_v = 0 \text{ for } v \in \Sigma\}$ and $\mathbb{Z}_\Sigma = \prod_{p \in \Sigma} \mathbb{Z}_p$, where p runs over finite places in Σ. We then define $\mathbb{Z}_{(\Sigma)} = \mathbb{Z}_\Sigma \cap \mathbb{Q}$. If $\Sigma = \{\infty\}$, $\{p\}$, or $\{p, \infty\}$ for a prime p, we write $\mathbb{A}^{(\infty)}$, $\mathbb{A}^{(p)}$ or $\mathbb{A}^{(p\infty)}$ for $\mathbb{A}^{(\Sigma)}$ accordingly. If Σ is a set of finite places, we write $\mathbb{A}^{(\Sigma\infty)}$ for $\mathbb{A}^{(\Sigma \cup \{\infty\})}$. We identify ${\mathbb{A}^{(\Sigma)}}^\times$ with $\{x \in \mathbb{A}^\times | x_v = 1 \text{ for } v \in \Sigma\}$ in an obvious way. The maximal compact subring of $\mathbb{A}^{(\infty)}$ is denoted by $\widehat{\mathbb{Z}}$, which is identified with the profinite ring $\prod_p \mathbb{Z}_p = \varprojlim_N \mathbb{Z}/N\mathbb{Z}$. For any $\mathbb{Z}$-module $\mathfrak{c}$ of finite type, we write $\widehat{\mathfrak{c}}$ for $\mathfrak{c} \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}$.

An algebraic group T (defined over a subring R of $\overline{\mathbb{Q}}$) is called a torus if its scalar extension $T_{/\overline{\mathbb{Q}}} = T \otimes_R \overline{\mathbb{Q}}$ is isomorphic to a product $\mathbb{G}_m^r$ of copies of the multiplicative group $\mathbb{G}_m$. The character group $X^*(T) = \mathrm{Hom}_{\text{alg-gp}}(T_{/\overline{\mathbb{Q}}}, \mathbb{G}_{m/\overline{\mathbb{Q}}})$ is simply denoted by $X(T)$, and elements of $X(T)$ are often called weights of T.

The symbol G is exclusively used for a reductive linear algebraic group defined over $\mathbb{Q}$. Then G_1 is its derived group of G (as an algebraic group defined over $\mathbb{Q}$), and Z is the center of G. We use the symbol $PG = G^{ad}$ for G/Z as an algebraic group over $\mathbb{Q}$. For the real points $G(\mathbb{R})$ of G, we regard it as a Lie group, and $G(\mathbb{R})^+$ is the identity-connected component of the Lie group $G(\mathbb{R})$. The symbol B often indicates a standard (upper triangular) Borel subgroup inside G, and the letters U and N are used to indicate the unipotent radical of B. The letter U is used when N is used to indicate an integer (and N is used when U is used to denote a unitary group). The torus B/U is denoted by T_G, and the maximal torus of G_1 induced by T_G is denoted by T.

2

Geometric Reciprocity Laws

Since reciprocity laws play an important role in this book, we now describe several reciprocity laws appearing in number theory in terms of automorphism groups of a field fixing a given prime or a given point. Only this chapter contains (elementary) exercises.

To fix our idea, we describe here briefly a model reciprocity law. For a given field $\mathfrak{K}$, we consider the field automorphism group $\mathrm{Aut}(\mathfrak{K})$ equipped with the Krull topology (which is described in Section 2.3). As is clear from the introduction, a geometric global reciprocity law gives a canonical description of $\mathrm{Aut}(\mathfrak{K})$ by the adelic points of an algebraic (reductive) group $G_{/\mathbb{Q}}$ (modulo rational center $Z(\mathbb{Q})$): $\mathrm{Aut}(\mathfrak{K}) \hookrightarrow \frac{G(\mathbb{A}^{(\infty)})}{Z(\mathbb{Q})}$ typically. If a local ring $\mathcal{V}$ with quotient field $\mathfrak{K}$ is given, we could define a decomposition subgroup D of $\mathrm{Aut}(\mathfrak{K})$ (equipped with the Krull topology) by

$$D = \left\{\sigma \in \mathrm{Aut}(\mathfrak{K}) \middle| \sigma(\mathcal{V}) = \mathcal{V}\right\},$$

and the local reciprocity law for $\mathcal{V}$ describes D via a well-defined algebraic subgroup H of G. Often D is given by the image of adele points, $H(\mathbb{A}^{(\infty)})$ (or $\mathbb{Q}_p$-points, $H(\mathbb{Q}_p)$), of H in $\frac{G(\mathbb{A}^{(\infty)})}{Z(\mathbb{Q})}$. This description may involve a subtle use of class field theory, and class field theory supplies us with one of the simplest examples of the reciprocity laws. The global reciprocity law is a particular case of the local one (taking $\mathcal{V} = \mathfrak{K}$). To have such reciprocity laws, the field $\mathfrak{K}$ cannot be finitely generated over its prime subfield except for the case where G is finite. So for an algebraic group G of positive dimension, the reciprocity laws get more geometric flavor, and $\mathfrak{K}$ is then realized as the function field of an infinite tower V of (geometrically irreducible) algebraic varieties defined over a tower of finite extensions over the prime field. Then if we have local reciprocity laws covering sufficiently many points of the tower V, the laws characterize a global model V of $\mathfrak{K}$ (a model of a field K is an algebraic variety whose function field is isomorphic to K in a canonical way). Thus to have a coherent system of local reciprocity laws is almost equivalent to having a well-defined global model of a given function field (of infinite type). Similarly, if V

is minimal, the algebrogeometric automorphism group $\mathrm{Aut}(V)$ of the variety V coincides with $\mathrm{Aut}(\mathfrak{K})$, and hence the tower is realized as a collection of algebraic varieties $\{V_S|V/S\}_S$ for S running through open compact subgroups S of $\mathrm{Aut}(\mathfrak{K})$, where V/S is the quotient variety of V by the action of S. Thus by the global reciprocity law, we can recover each member of the tower. This type of dichotomy appears in the theory of Shimura varieties.

In the later part of the book, we study more modern reciprocity laws (involving reductive groups of positive dimension) principally created by Shimura as an explicit description of automorphisms of Shimura varieties. In this beginning part of the book, we describe some reciprocity laws (including those relating finite groups G with the field automorphism groups) that only require minimal knowledge of algebraic geometry. Our requirement is reasonable knowledge of the theory of algebraic curves over number fields in this chapter. Some of the reciprocity laws are stated here without proof, though in the later chapters proofs are given in a more general setting. It would be a good exercise for the readers to deduce the results described in this chapter from the more elaborate versions in the later chapters. In the following chapter, we extend our scope in order to incorporate integral models of modular curves, and in the later chapters, we study Shimura varieties and their integral models via the language of schemes.

2.1 Sketch of Classical Reciprocity Laws

In this section, we sketch the classical reciprocity laws starting with the historic quadratic reciprocity invented by Euler (and proved by Gauss), and ending with a reciprocity law for a single rational elliptic curve. Here, primes p and q are always distinct odd primes.

2.1.1 Quadratic Reciprocity Law

For an integer n ($p \nmid n$), the Legendre symbol $\left(\frac{n}{p}\right)$ is defined by

$$\left(\frac{n}{p}\right) = \begin{cases} 1 & \text{if } x^2 \equiv n \mod p \text{ has a solution,} \\ -1 & \text{otherwise.} \end{cases}$$

Since $x^2 \equiv n \mod p$ has a solution if and only if $n \in (\mathbb{F}_p^\times)^2$ for $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$, $n \mapsto \left(\frac{n}{p}\right)$ gives an identification of $\mathbb{F}_p^\times/(\mathbb{F}_p^\times)^2$ with $\{\pm 1\}$; in particular, $n \mapsto \left(\frac{n}{p}\right)$ is a character of the finite multiplicative group $(\mathbb{Z}/p\mathbb{Z})^\times$.

The quadratic reciprocity law guessed by Euler and proven by Gauss,

$$\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4}\left(\frac{p}{q}\right) \quad \text{(Legendre, 1785),}$$

has an equivalent formulation due to Euler

$$\left(\frac{q}{p}\right) = \left(\frac{p^*}{q}\right) \quad \text{(Euler, 1744)},$$

where $p^* = (-1)^{(p-1)/2}p$. For a general odd positive integer n prime to p, define $\left(\frac{p^*}{n}\right) = \prod_q \left(\frac{p^*}{q}\right)^{e(q)}$ if the prime factorization of n is given by $\prod_q q^{e(q)}$. Thus $n \mapsto \left(\frac{p^*}{n}\right)$ is equal to the character: $n \mapsto \left(\frac{n}{p}\right)$.

This character decides how a prime decomposes in $\mathfrak{K} = \mathbb{Q}[\sqrt{p^*}]$. We look into the ring $R = \mathbb{Z}[\sqrt{p^*}] = \mathbb{Z}[X]/(X^2 - p^*)$. In R, a prime ideal (q) of $\mathbb{Z}$ could remain prime or become a product of two prime ideals; that is, $(q) = \mathfrak{q}\mathfrak{q}'$ or $(q) = \mathfrak{q}$ for prime ideals $\mathfrak{q}, \mathfrak{q}'$ in R. Note that $R/(q) = \mathbb{F}_q[X]/(X^2 - p^*)$. The polynomial $X^2 - p^*$ is reducible in $\mathbb{F}_q[X]$ if and only if it has a solution in $\mathbb{F}_q$:

$$(q) = \mathfrak{q}\mathfrak{q}' \iff R/(q) \cong \mathbb{F}_q \oplus \mathbb{F}_q \iff \left(\frac{p^*}{q}\right) = 1,$$

where $\mathfrak{q}$ is the kernel of the projection of R onto the first factor $\mathbb{F}_q$ of $\mathbb{F}_q \oplus \mathbb{F}_q$.

The nontrivial automorphism σ of $\mathbb{Q}[\sqrt{p^*}]$ interchanges the roots of $X^2 - p^*$ and interchanges $\mathfrak{q}$ and $\mathfrak{q}'$; so, $\mathfrak{q}' = \sigma(\mathfrak{q})$. Identifying $\{\pm 1\}$ with $\mathrm{Aut}(\mathfrak{K}) = \mathrm{Gal}(\mathbb{Q}[\sqrt{p^*}]/\mathbb{Q})$, $\left(\frac{p^*}{q}\right)$ gives the generator of the subgroup of $\mathrm{Aut}(\mathfrak{K})$ fixing a prime factor of (q).

If $(q) = \mathfrak{q}$ remains prime, $\dim_{\mathbb{F}_q} R/(q) = 2$, and $\sigma(\mathfrak{q}) = \mathfrak{q}$. The stabilizer of $\mathfrak{q}$ is the entire Galois group, and $\langle\left(\frac{p^*}{q}\right)\rangle = \{\pm 1\} \cong \mathrm{Aut}(\mathfrak{K})$ gives the stabilizer. In summary, identifying $\mathrm{Aut}(\mathfrak{K})$ with $\{\pm 1\}$, the subgroup generated by $\left(\frac{p^*}{q}\right)$ gives the stabilizer of a prime ideal $\mathfrak{q}|q$ in R (in other words, the stabilizer of the $\mathfrak{q}$-adic valuation ring $\mathcal{V}$ of $\mathfrak{K}$). The stabilizer is called the decomposition subgroup of q. The information of the decomposition group of q is equivalent to knowing how the prime (q) splits in R.

2.1.2 Cyclotomic Version

In the mid nineteenth century, Kummer extended the quadratic reciprocity law to cyclotomic fields (in his study of decomposition of prime numbers into a product of his "ideal" prime numbers). Let μ_p be the group of all pth roots of unity, and consider the extension $\mathfrak{K} = \mathbb{Q}[\mu_p]$ generated by pth roots of unity. Fixing one primitive root of unity, say, $\zeta = \zeta_p = \exp\left(\frac{2\pi\sqrt{-1}}{p}\right)$, μ_p is a cyclic group of order p generated by ζ_p. Each automorphism σ of μ_p takes ζ to another primitive root of unity ζ^m. Since ζ^m is primitive, $p \nmid m$ and hence, we have an identification $\mathrm{Aut}(\mu_p) \cong \mathbb{F}_p^\times$ by $\chi_p : \sigma \mapsto m$. Actually the finite flat group scheme μ_p over $\mathbb{Q}$ is the model of its function field $\mathfrak{K} = \mathbb{Q}[\mu_p]$, and $\mathrm{Aut}(\mathfrak{K}) = \mathrm{Gal}(\mathbb{Q}[\mu_p]/\mathbb{Q}) = \mathrm{Aut}(\mu_p)$. Since $\sigma \in \mathrm{Gal}(\mathbb{Q}[\mu_p]/\mathbb{Q})$ induces an

element in $\mathrm{Aut}(\mu_p)$ (and $[\mathbb{Q}[\mu_p]:\mathbb{Q}] = |\mathbb{F}_p^\times|$), we see that $\mathrm{Gal}(\mathbb{Q}[\mu_p]/\mathbb{Q}) \cong \mathbb{F}_p^\times$ by χ_p. We write $\phi_q \in \mathrm{Gal}(\mathbb{Q}[\mu_p]/\mathbb{Q})$ with $\chi_p(\phi_q) = q$; so, $\phi_q(\zeta) = \zeta^q$.

Again we ask how a prime (q) decomposes in the ring $R = \mathbb{Z}[\mu_p]$. Pick a prime ideal $\mathfrak{q}|q$ in R; we find that $R/\mathfrak{q}$ is a finite extension of $\mathbb{F}_q$; so, it is of the form $\mathbb{F}_{q^f}$ for $f = \dim_{\mathbb{F}_q} R/\mathfrak{q}$. The Galois group $\mathrm{Gal}(\mathbb{F}_{q^f}/\mathbb{F}_q)$ is a cyclic group of order f generated by a canonical generator F taking $x \in \mathbb{F}_{q^f}$ to $x^q \in \mathbb{F}_{q^f}$; thus, the automorphism F of $\mathbb{F}_{q^f}$ is induced by ϕ_q. In other words,

- *The decomposition group of* $\mathfrak{q}|q$ *is given by* $\langle\phi_q\rangle \cong \langle q\rangle \subset \mathbb{F}_p^\times$, *and*
- q *is of order* f *in* $\mathbb{F}_p^\times \iff (q) = \mathfrak{q}\sigma(\mathfrak{q})\cdots\sigma^{g-1}(\mathfrak{q})$ *in* R *for the integer* $g = [\mathbb{Q}[\mu_p]:\mathbb{Q}]/f$,

where σ is the generator of $\mathrm{Gal}(\mathbb{Q}[\mu_p]/\mathbb{Q})/\langle\phi_q\rangle$. Thus one feature of the reciprocity law is the determination of the decomposition group of a given prime q in a given Galois extension $K/\mathbb{Q}$.

2.1.3 Geometric Interpretation

To further generalize the law, we need to ponder a philosophical reason why we have such an arithmetic way of describing the decomposition group. One feature of the cyclotomic version is the existence of a canonical generator ϕ_q (the *Frobenius element* at q) of the decomposition group, and another is the appearance of the exponential function $\exp(z)$, which has the following fundamental identity,

$$\exp\left(2\pi i\frac{1}{p}\right)^{\phi_q} = \phi_q(\zeta) = \zeta^q = \exp\left(2\pi i\frac{q}{p}\right).$$

So, roughly, the complex analytic function $\exp : \mathbb{C} \to \mathbb{C}^\times$ contains all the information of the reciprocity law, and $\exp\left(2\pi i\frac{1}{p}\right)$ gives a canonical generator ζ of the field $\mathbb{Q}[\mu_p]$ and behaves nicely under its Galois conjugation.

The function exp gives rise to the following exact sequence,

$$0 \longrightarrow 2\pi i\mathbb{Z} \longrightarrow \mathbb{C} \xrightarrow{\exp} \mathbb{C}^\times \to 1,$$

and thus we evaluated the function exp at the fraction $2\pi i\frac{1}{p} \in \frac{1}{p}2\pi i\mathbb{Z}/2\pi i\mathbb{Z}$. The period $2\pi i$ of the exponential function "exp" is important to ensure that the value of "exp" on $2\pi i\mathbb{Q}$ is algebraic, and we have $2\pi i = \int_\gamma \frac{dt}{t}$ for the unit circle γ which generates the fundamental group $\pi_1(\mathbb{C}^\times) = H_1(\mathbb{C}^\times, \mathbb{Z})$.

Hilbert asked in his twelfth problem (of his famous lecture in 1900 at the International Congress of Mathematicians held in Paris; [Hl]), for a given Galois extension (actually an abelian extension in his original setting),

Is there any complex analytic function which describes fully the reciprocity law of the extension?

2.1.4 Kronecker's Reciprocity Law

In the language of Poincaré, the fundamental group of $\mathbb{C}^\times$ is given by $\mathbb{Z}$, and the universal covering of $\mathbb{C}^\times$ is given by $\mathbb{C} \xrightarrow{\exp} \mathbb{C}^\times$. We can create such a situation starting with an imaginary quadratic field $M = \mathbb{Q}[\sqrt{-D}]$ with discriminant $-D < 0$ in place of $\mathbb{Q}$. For any given ideal $\mathfrak{a} \subset R$ (for the integer ring $R \subset M$) and a period $\Omega \in \mathbb{C}^\times$, we consider an exact sequence:

$$0 \longrightarrow \Omega\mathfrak{a} \longrightarrow \mathbb{C} \xrightarrow{(\mathcal{P},\mathcal{P}')} E(\mathbb{C}) \longrightarrow 0.$$

For the moment $E(\mathbb{C})$ is a quotient space $\mathbb{C}/\mathfrak{a}$, which is a Riemann surface of genus 1. The period Ω of the elliptic curve E is inserted here to ensure that the value on ΩM of the Weierstrass functions $\mathcal{P}$ and $\mathcal{P}'$ (as defined below) are algebraic. Writing ω for the (translation-) invariant differential on $E(\mathbb{C})$ induced by du for the variable u of $\mathbb{C}$, we have $\Omega = \int_\gamma \omega$ for the generator γ of $\pi_1(E(\mathbb{C})) = H_1(E(\mathbb{C}), \mathbb{Z}) = \mathfrak{a}\gamma$ just as before. Indeed, for an $\mathbb{R}$-base (w_1, w_2) of $\mathbb{C}$, we can think of $E_L(\mathbb{C}) = \mathbb{C}/L$ for $L = \mathbb{Z}w_1 + \mathbb{Z}w_2$ replacing $\mathfrak{a}$ by L.

Weierstrass studied analysis and geometry of Riemann surfaces (transforming Riemann's marvelous but rather intuitive ideas into a rigorous mathematics). In particular, for the Riemann surface $E(\mathbb{C})$ of genus 1, he created the following function well-defined over $E(\mathbb{C})$,

$$x(u) = \mathcal{P}(u) = \frac{1}{u^2} + \sum_{\ell \in L - \{0\}} \left(\frac{1}{(u-\ell)^2} - \frac{1}{\ell^2} \right),$$

averaging the translations of $\frac{1}{u^2}$ over L, which is a two-dimensional analogue of the partial fraction expansion of the cotangent function (see [LFE] 2.1):

$$\pi \cot(\pi z) = \frac{1}{z} + \sum_{n=1}^{\infty} \left\{ \frac{1}{z+n} + \frac{1}{z-n} \right\}.$$

This function converges absolutely over $\mathbb{C} - L$ and gives a meromorphic function on $E(\mathbb{C})$. The Laurent expansion of $\mathcal{P}$ and its derivative can be easily computed, and we have

$$x(u) = \mathcal{P}(u) = \frac{1}{u^2} + \frac{g_2}{20}u^2 + \frac{g_3}{28}u^4 + \cdots, \quad y(u) = \mathcal{P}'(u) = -\frac{2}{u^3} + \sum_{n \geq 1} a_n u^n,$$

where $g_2 = g_2(L) = 60 \sum_{\ell \neq 0} \ell^{-4}$ and $g_3 = g_3(L) = 140 \sum_{\ell \neq 0} \ell^{-6}$. The constants g_2 and g_3 are actually complex analytic functions of $w = (w_1, w_2)$.

Canceling the poles, we consider $\varphi = y^2 - 4x^3 + g_2 x + g_3$ which is holomorphic everywhere on a compact Riemann surface $E(\mathbb{C})$; so, it has to be constant. The function φ has to be identically 0, because φ has no constant term. Thus $u \mapsto \mathbf{E}(u) = (u^3 x(u), u^3 y(u), u^3) \in \mathbf{P}^2$ embeds the Riemann surface into the projective space of dimension 2, whose image is an algebraic curve (called an *elliptic curve*) defined by the homogeneous equation:

$Y^2Z = 4X^3 - g_2XZ^2 - g_3Z^3$ ($x = X/Z$ and $y = Y/Z$). We write $\Delta = g_3^2 - 27g_2^3$ for the discriminant of $4x^3 - g_2x - g_3$.

If we pick $\lambda \in \mathbb{C}^\times$, we have a biholomorphic isomorphism of Riemann surfaces: $u \mapsto \lambda u$ of $E_L(\mathbb{C}) \cong E_{\lambda L}(\mathbb{C})$. By definition, $g_i(\lambda L) = \lambda^{-2i} g_i(L)$ and $\Delta(\lambda L) = \lambda^{-12}\Delta(L)$. We have $j(\lambda L) = j(L)$ for the *j-invariant* defined by $j(w) = j(L) = \frac{g_2^3}{\Delta}$. The following facts are known from the general theory of algebraic curves and are proven in this chapter.

- If $\phi : E_L \to E_{L'}$ is a holomorphic homomorphism with $\phi(0) = 0$, then there exists $\lambda \in \mathbb{C}^\times$ such that $L' \supset \lambda L$ and $\phi(u+L) = (\lambda u + L')$ (Theorem 2.39);
- Write a morphism $\phi : E_L \to E_{L'}$ as $\phi(u) = (\phi_x(u), \phi_y(u), 1)$ using the coordinates of the projective space $\mathbf{P}^2$. Then ϕ_x and ϕ_y are rational functions of x_L and y_L; that is, $\phi_x = \frac{A(x,y,g_2,g_3)}{B(x,y,g_2,g_3)}$ and $\phi_y = \frac{\alpha(x,y,g_2,g_3)}{\beta(x,y,g_2,g_3)}$ for polynomials A, B, α, β with coefficients in $\mathbb{Q}$ independent of L (Corollary 2.26);
- Every genus 1 Riemann surface is obtained as E_L for an L (Theorem 2.39);
- $E_{L/\mathbb{C}} \cong E_{L'/\mathbb{C}} \iff j(L) = j(L')$ (Corollary 2.35).

The function j is an example of a *modular function* and g_2 and g_3 are examples of *modular forms*. Modular forms are a special kind of automorphic forms defined on a more general complex domain.

For a given number field F, the maximal everywhere unramified abelian extension H/F is called the Hilbert class field. Kronecker studied the Hilbert class field H of the imaginary quadratic field M. Take a prime $\mathfrak{p}$ of the integer ring R of M and its prime factor $\mathfrak{P}$ in the integer ring O_H of H. The decomposition group $D_{\mathfrak{P}'} = \{\sigma \in \mathrm{Gal}(H/M) | \sigma(\mathfrak{P}') = \mathfrak{P}'\}$ for any other prime factor $\mathfrak{P}'|\mathfrak{p}$ is a conjugate of $D_{\mathfrak{P}}$ (as is well known). Since $\mathrm{Gal}(H/M)$ is abelian, the group $D_{\mathfrak{P}}$ is uniquely determined independently of the factor $\mathfrak{P}$ of $\mathfrak{p}$; so, we write $D_{\mathfrak{p}}$ for $D_{\mathfrak{P}}$. Since H/M is unramified (at $\mathfrak{p}$), $D_{\mathfrak{p}} \cong \mathrm{Gal}(K/k)$ for the residue fields $K = O_H/\mathfrak{P}$ and $k = R/\mathfrak{p} = \mathbb{F}_q$ for a prime power q. The Frobenius automorphism $\phi_{\mathfrak{p}} \in D_{\mathfrak{p}}$ sending $x \in K$ to $x^q \in K$ gives a canonical generator of the Galois group $\mathrm{Gal}(K/k)$. What Kronecker found is

Theorem (Kronecker–Weber) *Let H/M be the Hilbert class field of M. Then $H = M[j(\mathfrak{a})]$ for any ideal $0 \neq \mathfrak{a} \subset R$ and $j(\mathfrak{a})^{\phi_{\mathfrak{p}}} = j(\mathfrak{p}^{-1}\mathfrak{a})$. In particular, $\mathrm{Gal}(H/M)$ is isomorphic to the ideal class group of M.*

Proof. We give a sketch of a proof. We first explain why $j(\mathfrak{a})$ is an algebraic number. If $\mathrm{End}(E_L)$ contains the integer ring R of an imaginary quadratic field M, then $RL \subset L$; so, $L \subset \Omega M$ for $0 \neq \Omega \in L$. Via multiplication by Ω, $\mathbb{C}/\mathfrak{a} \cong \mathbb{C}/L$ for $\mathfrak{a} = \Omega^{-1}L$. Thus we may assume that L is an ideal $\mathfrak{a}$ of M. Since the isomorphism class of $E_{\mathfrak{a}}$ only depends on $\mathfrak{a}$ up to scalar multiplication, there are only countably many isomorphism classes of E_L with $\mathrm{End}(E_L) \supset R$. Regard $E_{\mathfrak{a}}$ as a curve defined by the equation $y^2 = 4x^3 - g_2x - g_3$, and consider its conjugate $E_{\mathfrak{a}}^\sigma$ defined by $y^2 = 4x^3 - \sigma(g_2)x - \sigma(g_3)$ for any field automorphism σ of $\mathbb{C}$. Then $\mathrm{End}(E_{\mathfrak{a}}^\sigma)$ contains R because all endomorphisms are rational functions of (x, y, g_2, g_3); so, applying σ to their coefficients, we

get an endomorphism $\phi^\sigma \in \mathrm{End}(E_{\mathfrak{a}}^\sigma)$ from $\phi \in \mathrm{End}(E_{\mathfrak{a}})$. The morphism $\phi \mapsto \phi^\sigma$ is an isomorphism of rings. Thus $E_{\mathfrak{a}}^\sigma = E_{\mathfrak{b}}$ for an ideal $\mathfrak{b}$ in M and $j(\mathfrak{a})^\sigma = j(\mathfrak{b})$; so, $j(\mathfrak{a})$ has only countably many conjugates. This implies that they are finitely many, because if $x \in \mathbb{C}$ is transcendental over $\mathbb{Q}$, one can embed $\mathbb{Q}(x)$ into $\mathbb{C}$ in continuously many different ways ($\mathbb{C}$ has continuously many transcendental numbers). Thus the number of conjugates of $j(\mathfrak{a})$ is finite, and hence $j(\mathfrak{a})$ is algebraic. Since $j(\mathfrak{b}) = j(\mathfrak{a})$ if and only if $\mathfrak{b} = \alpha\mathfrak{a}$ for some $\alpha \in M^\times$, the fractional ideals of M modulo scalar multiplication are finitely many. For a given j, we can create a Weierstrass equation $y^2 = 4x^3 - g_2x - g_3$ in $\mathbb{Q}(j)[x, y]$ with invariant j, for example, if $j \notin \{0, 1\}$, we put $g_2 = g_3 = \frac{27j}{j-1}$ (see just above Corollary 2.35 in the text or [IAT] 4.1). In other words, by choosing the period Ω as $\Omega = \int_\gamma \frac{dx}{y} = \int_\gamma \omega \in \mathbb{C}^\times$, we have

$$\Omega^{-4} g_2(\mathfrak{a}) = g_2(\Omega\mathfrak{a}) = \frac{27j}{j-1} = g_3(\Omega\mathfrak{a}) \in \overline{\mathbb{Q}}.$$

For almost all prime ideals $\mathfrak{P}$ of $\mathbb{Q}(j(\mathfrak{a}))$, g_2 and g_3 are $\mathfrak{P}$-integral with $\Delta(\mathfrak{a}) \not\equiv 0 \mod \mathfrak{P}$, and the equation modulo $\mathfrak{P}$ gives rise to an elliptic curve $\widetilde{E}_{\mathfrak{a}}$ defined over a finite field $\mathbb{F}$. The relative Frobenius map $F : (x, y) \mapsto (x^p, y^p)$ sends $\widetilde{E}_{\mathfrak{a}}$ to another elliptic curve $\widetilde{E}_{\mathfrak{a}}^{(p)}$ defined by $y^2 = 4x^3 - g_2^p x - g_3^p \in \mathbb{F}[x, y]$. Suppose that $(p) = \mathfrak{P} \cap \mathbb{Z}$ is split into $\mathfrak{p}\overline{\mathfrak{p}}$ in the imaginary quadratic field M. Then the group of p-torsion points $E_{\mathfrak{a}}[p]$ is the direct sum of the group of $\mathfrak{p}$-torsion points and that of $\overline{\mathfrak{p}}$: $E_{\mathfrak{a}}[p] = E_{\mathfrak{a}}[\mathfrak{p}] \oplus E_{\mathfrak{a}}[\overline{\mathfrak{p}}]$. As is well known (cf. [GME] 2.9.1), we can choose $\mathfrak{p}$ uniquely so that $\widetilde{E}_{\mathfrak{a}}[\mathfrak{p}] \cong R/\mathfrak{p}$ having p distinct points. Since $F : \widetilde{E}_{\mathfrak{a}} \to \widetilde{E}_{\mathfrak{a}}^{(p)}$ is a homomorphism of groups that is a zero map on the tangent space, it induces a group isomorphism $E_{\mathfrak{a}}(\overline{\mathbb{F}}) \cong E_{\mathfrak{a}}^{(p)}(\overline{\mathbb{F}})$ (though not a scheme isomorphism), and hence $\widetilde{E}_{\mathfrak{a}}^{(p)} \cong \widetilde{E}_{\mathfrak{p}^{-1}\mathfrak{a}}$, because $p : \widetilde{E}_{\mathfrak{a}} \to \widetilde{E}_{\mathfrak{a}}$ is an endomorphism (see [GME] 2.9.1) which factors $\widetilde{E}_{\mathfrak{a}} \xrightarrow{F} \widetilde{E}_{\mathfrak{a}}^{(p)} \xrightarrow{V} \widetilde{E}_{\mathfrak{a}}$ with $\mathrm{Ker}(V) = \widetilde{E}_{\mathfrak{a}}^{(p)}[\mathfrak{p}] = \mathfrak{p}^{-1}\mathfrak{a}/\mathfrak{a}$. This shows that $j(\mathfrak{a})^{\phi_{\mathfrak{p}}} = j(\mathfrak{p}^{-1}\mathfrak{a})$. Since $j(\mathfrak{a})$ characterizes the isomorphism class of $E_{\mathfrak{a}}$ (over $\overline{\mathbb{Q}}$), we know that the $j(\mathfrak{a})$s indexed by $\mathfrak{a} \in Cl_M$ for the ideal class group Cl_M of M are all distinct. Since in each class of Cl_M, we can find split $\mathfrak{p}$ outside a given finite set of primes of M (the Chebotarev density), we find that $\{j(\mathfrak{a})\}_{\mathfrak{a} \in Cl_M}$ are all conjugates of each other over M; so, they span the Hilbert class field H/M. This proves the above theorem (and finiteness of Cl_M).

By the theorem (or its proof), $\mathrm{Gal}(H/M) \cong Cl_M$. This is the explicit class field theory for the imaginary quadratic field M. This type of result is generalized by Shimura, Taniyama, and Weil to *CM fields*, using a quotient $\mathbb{C}^d/L$ for a lattice L of higher rank (theory of abelian varieties with complex multiplication; see Theorem 4.19 in the text, [IAT] Chapter 5, and [ACM]). A CM field means a totally imaginary quadratic extension of a totally real field.

2.1.5 Reciprocity Law for Elliptic Curves

We can slightly generalize the above construction. Let

$$E_L[p] = \{u \in E_L(\mathbb{C}) | pu = 0\} = \frac{1}{p}L/L \cong \mathbb{F}_p^2.$$

Note that $\mathbf{E}(0) = (0, -2, 0) \in \mathbf{P}^2(\mathbb{Q})$, and hence the point 0 of E_L does not move after applying $\sigma \in \mathrm{Aut}(\mathbb{C})$. Then $\sigma \in \mathrm{Aut}(\mathbb{C})$ brings $p : E_L \to E_L$ to $p : E_L^\sigma \to E_L^\sigma$ because they are rational functions of (x, y, g_2, g_3); so, if E_L is defined over a number field $M' = \mathbb{Q}[g_2(L), g_3(L)]$, $\sigma \in \mathrm{Aut}(\mathbb{C}/M')$ induces a linear automorphism of $E_L[p]$. Taking a base $\mathbf{w} = (\frac{w_1}{p}, \frac{w_2}{p})$ of $E_L[p]$, we find $\sigma(\mathbf{w}) = \mathbf{w}\rho(\sigma)$ for a 2×2 matrix $\rho(\sigma) \in GL_2(\mathbb{F}_p)$, and so, we can identify $\mathrm{Gal}(M'(E_L[p])/M')$ with a subgroup $\mathrm{Im}(\rho)$ of $GL_2(\mathbb{F}_p)$. Here $M'(E_L[p])$ is the field generated by $x(\frac{w_i}{p})$ and $y(\frac{w_i}{p})$ $(i = 1, 2)$ over M'. Thus identifying $E_L[p]$ with the column vector space $\mathbb{F}_p^2$ by ${}^t(a, b) \mapsto \mathbf{w}\,{}^t(a, b) \in E_L[p]$, we have

$$x(\rho(\sigma)v) = x(v)^\sigma \quad \text{and} \quad y(\rho(\sigma)v) = y(v)^\sigma.$$

This reciprocity law is still half baked, because we have not made explicit the form of $\rho(\phi_\mathfrak{q})$ for the canonical generator $\phi_\mathfrak{q}$ of the decomposition group of a prime $\mathfrak{q}$ of M'. This can be done when L is a fractional ideal of an imaginary quadratic field, and the refined version is an example of Shimura's reciprocity laws (although the origin of this reciprocity goes back to Kronecker). When L is not in an imaginary quadratic field, $\mathrm{Im}(\rho)$ is almost always full (by a result of Serre; see [ARE] and [Se2]), and we can (conjecturally for general $M' \neq \mathbb{Q}$) make explicit $\mathrm{Tr}(\rho(\phi_\mathfrak{q}))$ and $\det(\rho(\phi_\mathfrak{q}))$ (see [GME] 5.2.4). However this information is still a bit short of completely determining the splitting of a prime $\mathfrak{q}$ in $M'(E_L[p])$ (see [Sh2]), and the analysis of the Galois representation $\rho : \mathrm{Gal}(M'(E_L[p])/M') \hookrightarrow GL_2(\mathbb{F}_p)$ is still a central subject today.

From what I said, it is clear that *the study of modular functions and modular forms (and their generalization, often called automorphic forms) is natural and crucial in algebraic number theory,* though they appear to be rather analytic and geometric objects.

2.2 Cyclotomic Reciprocity Laws and Adeles

We now give a more detailed description of the cyclotomic reciprocity laws and relate them to the idele class group of $\mathbb{Q}$. This fact was found basically by Kronecker and is one of the simplest examples of the reciprocity laws in class field theory.

2.2.1 Cyclotomic Fields

We look into the exponential function $\mathbf{e} : \mathbb{C} \twoheadrightarrow \mathbb{C}^\times$ given by $\mathbf{e}(z) = \exp(2\pi i z)$. For a positive integer N, we write $\zeta = \zeta_N = \mathbf{e}\left(\frac{1}{N}\right)$. Then ζ is a primitive

Nth root of unity. When N is a prime p, it satisfies the equation $X^p - 1 = (X-1)\Phi_p(X) = 0$ for $\Phi_p(X) = X^{p-1} + X^{p-2} + \cdots + 1$. Since $\zeta \neq 1$, $\Phi_p(\zeta) = 0$.

Proposition 2.1 *The cyclotomic polynomial $\Phi_p(X)$ is irreducible in $\mathbb{Q}[X]$.*

Proof. The irreducibility of Φ_p and $\Psi(X) = \Phi_p(X+1)$ is equivalent. We see easily that the constant term is given by $\Psi(0) = \Phi_p(1) = p$. Since $(X+1)^p - 1 = X\Psi(X) \equiv X^p \mod p$, by the Eisenstein criterion, $\Psi(X)$ is irreducible. □

Corollary 2.2 *The equation $\Phi_{p^n}(X) = \Phi_p(X^{p^{n-1}})$ for an integer $n \geq 1$ is irreducible in $\mathbb{Q}[X]$.*

The proof of this corollary is left to the reader as an exercise.

Corollary 2.3 *Let $N = p^n$ for a prime p. The cyclotomic field $\mathbb{Q}[\zeta_N]$ is a Galois extension of $\mathbb{Q}$ whose Galois group is abelian. We have a canonical isomorphism of groups $\chi = \chi_N : \mathrm{Gal}(\mathbb{Q}[\zeta_N]/\mathbb{Q}) \cong (\mathbb{Z}/N\mathbb{Z})^\times$ given by $\sigma(\mathbf{e}(\frac{1}{N})) = \mathbf{e}\left(\frac{\chi(\sigma)}{N}\right)$ for $\sigma \in \mathrm{Gal}(\mathbb{Q}[\zeta_N]/\mathbb{Q})$.*

Proof. Let μ_N be the group of Nth roots of unity. Then $\mathbf{e}$ induces an isomorphism $\iota : \mathbb{Z}/N\mathbb{Z} \cong \mu_N$ of groups by $m \mapsto \mathbf{e}(\frac{m}{N})$. Since $\sigma(\zeta_N)$ is another primitive Nth root for any $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ ($\overline{\mathbb{Q}}$ is the algebraic closure of $\mathbb{Q}$ in $\mathbb{C}$), σ induces an automorphism of the field $\mathbb{Q}[\zeta_N]$ because $\sigma(\zeta_N)$ is again a power of ζ_N. Thus $\mathbb{Q}[\zeta_N]/\mathbb{Q}$ is a Galois extension. Since $\mathrm{Aut}(\mathbb{Z}/N\mathbb{Z}) = (\mathbb{Z}/N\mathbb{Z})^\times$ (multiplicative group) by $\phi \mapsto \phi(1)$, we find $\chi = \chi_N : \mathrm{Gal}(\mathbb{Q}[\zeta_N]/\mathbb{Q}) \to (\mathbb{Z}/N\mathbb{Z})^\times$. If $\sigma \in \mathrm{Ker}(\chi)$, σ fixes all Nth roots of unity; so, χ is injective. Its p-component χ_{p^n} has to be surjective, since

$$|\mathrm{Gal}(\mathbb{Q}[\zeta_{p^n}]/\mathbb{Q})| = [\mathbb{Q}[\zeta_{p^n}] : \mathbb{Q}] = \deg(\Phi_{p^n}) = p^{n-1}(p-1) = |(\mathbb{Z}/p^n\mathbb{Z})^\times|. \ \square$$

We show later the surjectivity of χ_N for general N (see Theorem 2.8).

Corollary 2.4 *Let $\mathbb{Z}_{(p)} = \left\{\frac{a}{b} \middle| a, b \in \mathbb{Z},\ b\mathbb{Z} + p\mathbb{Z} = \mathbb{Z}\right\}$. Then the integral domain $\mathcal{V} = \mathbb{Z}_{(p)}[X]/(\Phi_{p^n}(X))$ is a discrete valuation ring whose maximal ideal is generated by $\varpi = \zeta_{p^n} - 1$, where ζ_{p^n} is the image of X in $\mathcal{V}$ (so, it is a primitive p^nth root of unity which can be identified with $\mathbf{e}\left(\frac{1}{p^n}\right)$). The valuation ring $\mathcal{V}$ is fully ramified over $\mathbb{Z}_{(p)}$, that is, $(\varpi)^{[\mathcal{V}:\mathbb{Z}_{(p)}]} = (p)$.*

Proof. By the fundamental theorem of arithmetic, any $x = \frac{a}{b}$ can be written as $p^{v(x)}y$ for $y = \frac{d}{c}$ with $p \nmid c$ and $p \nmid d$; so, $y^{-1} \in \mathbb{Z}_{(p)}$. Thus $v : \mathbb{Q} \to \mathbb{Z} \cup \{\infty\}$ gives the valuation of $\mathbb{Q}$ and $\mathbb{Z}_{(p)} = \{x \in \mathbb{Q} | v(x) \geq 0\}$; so, $\mathbb{Z}_{(p)}$ is a discrete valuation ring of $\mathbb{Q}$.

Let $\varpi = \zeta_p - 1$, and let $\mathcal{V} = \mathbb{Z}_{(p)}[X]/(\Phi_p(X))$. Since $\Psi(X) \equiv X^{p-1} \mod p$, we see that $\mathcal{V}/(p) \cong \mathbb{F}_p[X]/(X^{p-1})$. Then

$$(1) \supset (X) \supset (X^2) \supset \cdots \supset (X^{p-1}) = (0)$$

are the only ideals of $\mathcal{V}/(p)$, because $(a_\jmath X^\jmath + a_{\jmath+1}X^{\jmath+1} + \cdots + a_{p-1}X^{p-1}) = (X)^\jmath$ if $a_\jmath \neq 0$ in $\mathbb{F}_p$. Since $p = \Psi(0) = \prod_{\sigma\in \mathrm{Gal}(\mathbb{Q}[\zeta_p]/\mathbb{Q})} \sigma(\varpi)$, we find that $(\varpi) \supset (p)$. Then by the homomorphism theorem applied to $\mathcal{V} \twoheadrightarrow \mathcal{V}/(p)$, only ideals between (p) and (ϖ) are $(\varpi)^m$ for $m = 1, 2, \ldots, p-1$. By induction on n, one can show that

$$(1) \supset (X) \supset (X^2) \supset \cdots \supset (X^{n(p-1)}) = (0)$$

are the only ideals of $\mathcal{V}/(p^n)$. For a given ideal $\mathfrak{a} \subset \mathcal{V}$, $\mathfrak{a} \cap \mathbb{Z}_{(p)} = (p^\ell)$ because $\mathbb{Z}_{(p)}$ is a discrete valuation ring (DVR). Thus $\mathfrak{a} = (\varpi)^m$ for $0 \leq m \leq \ell(p-1)$. This shows that $\mathcal{V}$ is a DVR with valuation w given by $w(x) = n \iff (x) = (\varpi)^n$. The same argument as above works well for $\mathcal{V} = \mathbb{Z}_{(p)}[X]/(\Phi_{p^n}(X))$ for $\varpi = \zeta_{p^n} - 1$, since $\Phi_{p^n}(X+1) \equiv X^{p^{n-1}(p-1)} \mod p$. □

Let $\mathcal{V}$ be a DVR. Then we can extend the valuation v of $\mathcal{V}$ to its field of fractions K by $v(\frac{a}{b}) = v(a) - v(b)$. In other words, $\mathcal{V} = \{x \in K | v(x) \geq 0\}$. For any $x \in K$, we have either $x \in \mathcal{V}$ or $x^{-1} \in \mathcal{V}$. In particular, if $x \in K$ is integral over $\mathcal{V}$, then x satisfies an equation

$$x^n + a_1 x^{n-1} + \cdots + a_n = 0 \iff 1 = -a_1 x^{-1} - a_2 x^{-2} - \cdots - a_n x^{-n}$$

with $a_\jmath \in \mathcal{V}$. If $x \notin \mathcal{V}$, $v(x) < 0 \iff v(x^{-1}) > 0$, and

$$0 = v(1) = v(-a_1 x^{-1} - a_2 x^{-2} - \cdots - a_n x^{-n}) \geq \min_\jmath (v(a_\jmath) + \jmath v(x^{-1})) > 0,$$

which is a contradiction; so, $x \in \mathcal{V}$, and $\mathcal{V}$ is integrally closed. In particular $\mathbb{Z}_{(p)}[\zeta_{p^n}]$ is the integral closure of $\mathbb{Z}_{(p)}$, and (p) fully ramifies in $\mathbb{Z}_{(p)}[\zeta_{p^n}]$.

2.2.2 Cyclotomic Reciprocity Laws

Let q be a prime different from p. We now look into $\mathcal{V}_q = \mathbb{Z}_{(q)}[X]/(\Phi_{p^n}(X)) = \mathbb{Z}_{(q)}[\zeta_{p^n}]$. We look into $\overline{\Phi}_{p^n}(X) = (\Phi_{p^n}(X) \mod q) \in \mathbb{F}_q[X]$. Then $\mathcal{V}_q/q\mathcal{V}_q = \mathbb{F}_q[X]/(\overline{\Phi}_{p^n}(X))$ and $\dim_{\mathbb{F}_q} \mathcal{V}_q/q\mathcal{V}_q = \mathrm{rank}_{\mathbb{Z}_{(q)}} \mathcal{V}_q = \deg(\Phi_{p^n}(X))$. Since we have $p^{n-1}(p-1)$ distinct primitive p^nth roots of unity in $\overline{\mathbb{F}}_q$, $\overline{\Phi}_{p^n}(X)$ does not have multiple roots, and we see $\mathcal{V}_q/q\mathcal{V}_q \cong \mathbb{F}_{q^{f_1}} \oplus \mathbb{F}_{q^{f_2}} \oplus \cdots \oplus \mathbb{F}_{q^{f_g}}$. The image of X in $\mathbb{F}_{q^{f_\jmath}}$ is a primitive p^nth root $\alpha = \alpha_\jmath$ of unity in $\mathbb{F}_{q^{f_\jmath}}$, and $\mathbb{F}_{q^{f_\jmath}} = \mathbb{F}_q[\alpha]$. Thus $f = f_1 = f_2 = \cdots = f_g$ is the minimal exponent so that $\mathbb{F}_{q^f}$ contains a primitive p^nth root of unity. Since $\mathbb{F}_{q^f}^\times$ is made up of $(q^f - 1)$th roots of unity, f is the minimal exponent such that $p^n | q^f - 1$. The ideal $\mathfrak{q}_\jmath = \mathrm{Ker}(\pi_\jmath)$ for each projection $\pi_\jmath : \mathcal{V}_q \twoheadrightarrow \mathcal{V}_q/q\mathcal{V}_q \twoheadrightarrow \mathbb{F}_{q^{f_\jmath}}$ is a prime of $\mathcal{V}_q$, and we have

$$(q) = \prod_{\jmath=1}^{g} \mathfrak{q}_\jmath \quad \text{and} \quad fg = [\mathcal{V}_q/q\mathcal{V}_q : \mathbb{F}_q] = \deg(\Phi_{p^n}) = p^{n-1}(p-1).$$

Summing up, we get

Proposition 2.5 *Each prime $q \neq p$ decomposes into a product $(q) = \prod_{j=1}^{g} \mathfrak{q}_j$ of prime ideals $\mathfrak{q}_j$ with $\mathcal{V}_q/\mathfrak{q}_j \cong \mathbb{F}_{q^f}$ for the minimal exponent f with $p^n | q^f - 1$. We also have $fg = p^{n-1}(p-1)$.*

Let $\mathfrak{q} = \mathfrak{q}_j$. Then $\mathcal{V}_q/\mathfrak{q}\mathcal{V}_q = \mathbb{F}_{q^f}$. Since $\mathrm{Gal}(\mathbb{F}_{q^f}/\mathbb{F}_q)$ is a cyclic group of order f generated by ϕ_q given by $\phi_q(x) = x^q$, it is induced by an element $\phi_q \in \mathrm{Gal}(\mathbb{Q}[\zeta_{p^n}]/\mathbb{Q})$ with $\chi_{p^n}(\phi_q) = q$. In other words, the stabilizer of $\mathfrak{q}$ is given by $D_q = \langle \phi_q \rangle \subset G = \mathrm{Gal}(\mathbb{Q}[\zeta_{p^n}]/\mathbb{Q})$ generated by ϕ_q. We have an injection $G/D_q \hookrightarrow \{\mathfrak{q}_1, \ldots, \mathfrak{q}_g\}$ by $\sigma D_q \mapsto \sigma(\mathfrak{q})$. Since $|D_q| = |\mathrm{Gal}(\mathbb{F}_{q^f}/\mathbb{F}_q)| = f$, the image has $[\mathbb{Q}[\zeta_{p^n}] : \mathbb{Q}]/f = g$ elements; so, we get

$$\mathrm{Gal}(\mathbb{Q}[\zeta_{p^n}]/\mathbb{Q})/D_q \cong \{\mathfrak{q}_1, \ldots, \mathfrak{q}_g\}.$$

We have proven:

Theorem 2.6 (Cyclotomic Reciprocity) *If $q \neq p$ is a prime, then for each prime factor $\mathfrak{q}$ in $\mathbb{Z}_{(q)}[\zeta_{p^n}]$, the decomposition subgroup of $\mathfrak{q}$ in $G = \mathrm{Gal}(\mathbb{Q}[\zeta_{p^n}]/\mathbb{Q})$: $D_\mathfrak{q} = \{\sigma \in G | \sigma(\mathfrak{q}) = \mathfrak{q}\}$ is given by the cyclic subgroup $\langle \phi_q \rangle$ generated by the Frobenius element ϕ_q with $\chi_{p^n}(\phi_q) = q$. In particular, $(q) = \prod_{\sigma \in G/D_\mathfrak{q}} \sigma(\mathfrak{q})$ and $\mathcal{V}_q/\mathfrak{q} \cong \mathbb{F}_{q^f}$ for the order f of q in $(\mathbb{Z}/p^n\mathbb{Z})^\times$.*

Let $\mathcal{V}_\mathfrak{q}$ be the localization of $\mathcal{V}_q$ at $\mathfrak{q}$. Since $\prod_j \mathfrak{q}_j = (q)$ in $\mathcal{V}_q$, we find that $\mathfrak{q}\mathcal{V}_\mathfrak{q} = (q)$, which is a unique maximal ideal. Thus $\mathfrak{q}^m/\mathfrak{q}^{m+1} \cong \mathcal{V}_\mathfrak{q}/\mathfrak{q} = \mathbb{F}_{q^f}$ is a field; so, there is no ideal between $\mathfrak{q}^m$ and $\mathfrak{q}^{m+1}$ in $\mathcal{V}_\mathfrak{q}$. If $\mathfrak{a}$ is an ideal of $\mathcal{V}_\mathfrak{q}$, $\mathfrak{a} \cap \mathbb{Z}_{(q)} = q^m$; so, $\mathfrak{a} = \mathfrak{q}^m$. In other words, all ideals of $\mathcal{V}_\mathfrak{q}$ are given by

$$\mathcal{V}_\mathfrak{q} \supset \mathfrak{q} = (q) \supset \mathfrak{q}^2 = (q^2) \supset \mathfrak{q}^3 \supset \cdots \supset (0).$$

Therefore $\mathcal{V}_\mathfrak{q}$ is a discrete valuation ring. We have $\mathcal{V}_q \subset \cap_{\mathfrak{q}|q} \mathcal{V}_\mathfrak{q}$. If $\frac{b}{a} \in \mathbb{Q}[\zeta_{p^n}]$ for $a, b \in \mathcal{V}_q$ is in the intersection, then a is prime to $\mathfrak{q}$ for all $\mathfrak{q}|q$; thus, a is prime to q. This implies $\mathcal{V}_q = \cap_{\mathfrak{q}|q} \mathcal{V}_\mathfrak{q}$, and $\mathcal{V}_q$ is integrally closed (Exercise 3).

Since $\zeta = \zeta_{p^n}$ is defined by the relation

$$1 + \zeta^{p^{n-1}} + \zeta^{2p^{n-1}} + \cdots + \zeta^{(p-1)p^{n-1}} = 0,$$

the $p^{n-1}(p-1)$ elements $1, \zeta, \ldots, \zeta^{p^{n-1}(p-1)-1}$ form a base of $\mathbb{Z}[\zeta_{p^n}]$ over $\mathbb{Z}$. Then we have

$$\bigcap_\mathfrak{q} \mathcal{V}_\mathfrak{q} = \bigcap_q \mathcal{V}_q = (\cap_q \mathbb{Z}_{(q)})1 + (\cap_q \mathbb{Z}_{(q)})\zeta + \cdots + (\cap_q \mathbb{Z}_{(q)})\zeta^{p^{n-1}(p-1)-1} = \mathbb{Z}[\zeta_{p^n}],$$

where $\mathfrak{q}$ runs over all prime ideals of $O = \mathbb{Z}[\zeta_{p^n}]$ and q runs over all rational primes. This shows that $\mathbb{Z}[\zeta_{p^n}]$ is the integer ring of $\mathbb{Q}[\zeta_{p^n}]$; in other words, $\mathbb{Z}[\zeta_{p^n}]$ is the integral closure of $\mathbb{Z}$ in $\mathbb{Q}[\zeta_{p^n}]$. Thus $O_\mathfrak{q} = \mathcal{V}_\mathfrak{q}$ for a prime ideal $\mathfrak{q}$ of O, and we can restate the reciprocity law as follows:

Corollary 2.7 *The integer ring O of $\mathbb{Q}[\zeta_{p^n}]$ is generated by ζ_{p^n} over $\mathbb{Z}$. For a prime $q \neq p$ and a prime factor $\mathfrak{q}|q$, the decomposition group $D_\mathfrak{q}$ is independent*

of the choice of $\mathfrak{q}$ *and generated by the Frobenius element* ϕ_q *with* $\chi_{p^n}(\phi_q) = q$. *We have* $(q) = \prod_{\sigma \in G/D_\mathfrak{q}} \sigma(\mathfrak{q})$ *and* $O/\mathfrak{q} = \mathbb{F}_{q^f}$ *for the order* f *of* q *in* $(\mathbb{Z}/p^n\mathbb{Z})^\times$. *The prime* p *is totally ramified in* O, *and the prime factor* $\mathfrak{p}$ *of* p *is principal and is generated by* $\zeta_{p^n} - 1$. *In particular,* $I_\mathfrak{p} = D_\mathfrak{p} = \mathrm{Gal}(\mathbb{Q}[\zeta_{p^n}]/\mathbb{Q})$ *for the inertia subgroup* $I_\mathfrak{p} \subset D_\mathfrak{p}$.

2.2.3 Adelic Reformulation

Recall that p and q are distinct primes, and put $N = p^m q^n$. Then we have a commutative diagram

$$\begin{array}{ccc} \mathbb{Q}[\zeta_N] & \longleftarrow & \mathbb{Q}[\zeta_{p^m}] \\ \uparrow & & \uparrow{\scriptstyle I_p} \\ \mathbb{Q}[\zeta_{q^n}] & \underset{I_q}{\longleftarrow} & \mathbb{Q}, \end{array}$$

where I_p denotes the inertia group at p. Since p is unramified in $\mathbb{Q}[\zeta_{q^n}]$, we find $I_p \cap I_q = \{1\}$. Since $I_p = \mathrm{Gal}(\mathbb{Q}[\zeta_{p^m}]/\mathbb{Q})$ and $\mathrm{Gal}(\mathbb{Q}[\zeta_N]/\mathbb{Q})$ is an abelian group (Corollary 2.3), we find that $\mathrm{Gal}(\mathbb{Q}[\zeta_N]/\mathbb{Q}) = I_p \times I_q \cong (\mathbb{Z}/N\mathbb{Z})^\times$ by χ_N. Repeating this process for $N = \prod_p p^{e(p)}$, we find

Theorem 2.8 *The cyclotomic character* χ_N *induces an isomorphism*

$$\mathrm{Gal}(\mathbb{Q}[\zeta_N]/\mathbb{Q}) \cong (\mathbb{Z}/N\mathbb{Z})^\times$$

of groups. The decomposition group of a prime $q \nmid N$ *is cyclic and generated by the Frobenius element* ϕ_q *with* $\chi_N(\phi_q) = q \in (\mathbb{Z}/N\mathbb{Z})^\times$. *For a prime* $p|N$, *writing* $N = p^e N'$ *with* $p \nmid N'$, *the inertia group* I_p *of* p *is given by* $\chi_N^{-1}((\mathbb{Z}/p^e\mathbb{Z})^\times)$, *identifying* $(\mathbb{Z}/N\mathbb{Z})^\times = (\mathbb{Z}/p^e\mathbb{Z})^\times \times (\mathbb{Z}/N'\mathbb{Z})^\times$. *The decomposition group for* $p|N$ *is given by* $I_p \times \langle \phi_p \rangle$, *where* $\chi_{N'}(\phi_p) = p \in (\mathbb{Z}/N'\mathbb{Z})^\times$.

By definition, if $M|N$, $\mathbb{Q}[\zeta_M] \subset \mathbb{Q}[\zeta_N]$ because $\zeta_M = \zeta_N^{N/M}$. Then it is easy to check that the following diagram commutes:

$$\begin{array}{ccc} \mathrm{Gal}(\mathbb{Q}[\zeta_N]/\mathbb{Q}) & \xrightarrow{\text{restriction}} & \mathrm{Gal}(\mathbb{Q}[\zeta_M]/\mathbb{Q}) \\ {\scriptstyle \chi_N}\downarrow & & \downarrow{\scriptstyle \chi_M} \\ (\mathbb{Z}/N\mathbb{Z})^\times & \xrightarrow[(x \mod N) \mapsto (x \mod M)]{} & (\mathbb{Z}/M\mathbb{Z})^\times. \end{array} \tag{2.1}$$

Thus the composite (inside $\overline{\mathbb{Q}}$) of $\mathbb{Q}[\zeta_N]$ for all positive integers N is actually a union $\mathbb{Q}^{cyc} = \bigcup_N \mathbb{Q}[\zeta_N]$.

Taking the projective limit, we have

$$\chi : \mathrm{Gal}(\mathbb{Q}^{cyc}/\mathbb{Q}) \cong \widehat{\mathbb{Z}}^\times = \varprojlim_N (\mathbb{Z}/N\mathbb{Z})^\times. \tag{2.2}$$

We compute $\widehat{\mathbb{Z}} = \varprojlim_N (\mathbb{Z}/N\mathbb{Z})$. We recall that the p-adic integer ring $\mathbb{Z}_p = \varprojlim_n (\mathbb{Z}/p^n\mathbb{Z})$ is formed by the formal series $\sum_{n \geq 0} c_n p^n$ with $0 \leq c_n \leq p-1$,

which is the valuation ring (and is the completion of $\mathbb{Z}_{(p)}$ with respect to the norm $|x|_p = p^{-v(x)}$). Since any integer can be expanded as above by p-adic expansion, $\mathbb{Z} \subset \mathbb{Z}_p$, and this inclusion is compatible with $\mathbb{Z}_{(p)} \subset \mathbb{Z}_p$. Writing the prime decomposition of N as $N = \prod_{p|N} p^{e(p)}$, we find by the Chinese remainder theorem

$$\widehat{\mathbb{Z}} = \varprojlim_N (\mathbb{Z}/N\mathbb{Z}) = \varprojlim_N \prod_{p|N} (\mathbb{Z}/p^{e(p)}\mathbb{Z}) = \prod_p \varprojlim_e (\mathbb{Z}/p^e\mathbb{Z}) = \prod_p \mathbb{Z}_p, \quad (2.3)$$

where p runs over all prime numbers in the last two products. We put the p-adic topology on $\mathbb{Z}_p$ so that a system of neighborhoods of $x \in \mathbb{Z}_p$ is given by $\{x + p^n\mathbb{Z}_p\}_{n\geq 0}$. This makes $\mathbb{Z}_p$ (resp. $\mathbb{Z}_p^\times$) a compact profinite ring (resp. group). We equip $\widehat{\mathbb{Z}}$ and $\widehat{\mathbb{Z}}^\times$ with the product topology of each component ($\mathbb{Z}_p$ and $\mathbb{Z}_p^\times$). Then $\widehat{\mathbb{Z}}$ (resp. $\widehat{\mathbb{Z}}^\times$) is a compact profinite ring (resp. group).

We now want to prove: $\mathrm{Gal}(\mathbb{Q}^{cyc}/\mathbb{Q}) \cong \mathbb{A}^\times/\mathbb{Q}^\times\mathbb{R}_+^\times = GL_1(\mathbb{A}^{(\infty)})/\mathbb{Q}_+^\times$ for the adele ring $\mathbb{A}$ using the above expression of $\mathrm{Gal}(\mathbb{Q}^{cyc}/\mathbb{Q})$. The adele ring $\mathbb{A}$ is a rather complicated ring, but we find it very useful later, and it contains all arithmetic information of the field of rational numbers $\mathbb{Q}$.

Let us recall the definition of $\mathbb{A}$. We consider the product ring $\mathbb{R} \times \prod_p \mathbb{Q}_p$, where p runs over all positive prime numbers, and $\mathbb{Q}_p$ is the field of fractions of $\mathbb{Z}_p$, and $\mathbb{Q} \subset \mathbb{Q}_p$, which is the completion of $\mathbb{Q}$ under the p-adic norm $|\ |_p$ as above. Each p-adic number $x \in \mathbb{Q}_p$ has an expansion $\sum_{n \gg -\infty} c_n p^n$ ($c_n = 0$ if n is very negative). Since $\mathbb{Z}_p^\times$ is made up of $\sum_{n=0}^\infty c_n p^n$ with $p \nmid c_0$, $\mathbb{Q}_p = \bigsqcup_n p^n\mathbb{Z}_p^\times \sqcup \{0\}$, where n runs over all integers. Regard $\mathbb{Q}$ inside $\mathbb{R} \times \prod_p \mathbb{Q}_p$ by sending $\xi \in \mathbb{Q}$ to $(\xi, \xi, \cdots, \overset{p}{\xi}, \cdots) \in \mathbb{R} \times \prod_p \mathbb{Q}_p$. Let $\mathbb{A}$ be the subring of $\mathbb{R} \times \prod_p \mathbb{Q}_p$ generated by $\mathbb{Q}$, $\widehat{\mathbb{Z}}$, and $\mathbb{R}$. We often write $x = (x_\infty, \ldots, x_p, \ldots)$ for an element $x \in \mathbb{R} \times \prod_p \mathbb{Q}_p$; so, $x_\infty \in \mathbb{R}$.

Proposition 2.9 *We have*

$$\mathbb{A} = \mathbb{A}' := \left\{ (x_p)_{p:prime,\infty} \in \mathbb{R} \times \prod_p \mathbb{Q}_p \middle| x_p \in \mathbb{Z}_p \text{ for almost all } p \right\}$$

and $\mathbb{A} = (\mathbb{R} \times \widehat{\mathbb{Z}}) + \mathbb{Q}$. *Here "almost all p" means "except for finitely many p."*

Proof. The ring $\mathbb{A}'$ is a subring of $\mathbb{R} \times \prod_p \mathbb{Q}_p$. For $\xi = \frac{a}{b} \in \mathbb{Q}$ $(a, b \in \mathbb{Z})$, if $p \nmid b$, we have $\xi \in \mathbb{Z}_p$. Thus $\xi \in \mathbb{A}'$, and $\mathbb{A} \subset \mathbb{A}'$. If $x \in \mathbb{A}'$, we expand $x_p = \sum_n c_n p^n$, and define $[x_p] = \sum_{n<0} c_n p^n$, which is a fraction with a p-power denominator and is called the fraction p-part of x_p. Then $[x_p] = 0$ for almost all p, and $[x] = \sum_p [x_p] \in \mathbb{Q}$. Then we look into $x_p - [x] = x_p - [x_p] - \sum_{q\neq p}[x_q] \in \mathbb{Q}_p$. Since the denominator of $[x_q]$ is prime to p, we find $[x_q] \in \mathbb{Z}_p$. Thus $x_p - [x] = x_p - [x_p] - \sum_{q\neq p}[x_q] \in \mathbb{Z}_p$, and hence $x - [x] \in \widehat{\mathbb{Z}}$ inside $\mathbb{A}'$. In particular, $x = (x - [x]) + [x] \in (\mathbb{R} \times \widehat{\mathbb{Z}}) + \mathbb{Q} \subset \mathbb{A}$. This shows the last identity. □

Writing the finite part $\mathbb{A}^{(\infty)} = \mathbb{A} \cap \prod_p \mathbb{Q}_p$, we thus have $\mathbb{A}^{(\infty)} = \widehat{\mathbb{Z}} \otimes_{\mathbb{Z}} \mathbb{Q}$.

Proposition 2.10 *We have* $\mathbb{A}^\times = \mathbb{Q}^\times(\mathbb{R}_+^\times \times \widehat{\mathbb{Z}}^\times)$ *by* $x \mapsto \langle x \rangle (x/\langle x \rangle)$, *where* $\mathbb{R}_+^\times = \{x \in \mathbb{R} | x > 0\}$.

Proof. Since $\mathbb{Q}_p = \bigsqcup_n p^n \mathbb{Z}_p^\times \sqcup \{0\}$, we can write $x_p \in \mathbb{Q}_p$ as $x_p = p^{v(x_p)} u_p$ with $u_p \in \mathbb{Z}_p^\times$. Then for $x \in \mathbb{A}^\times$, $x^{-1} \in \mathbb{A}$, and hence $|x_p|_p = 1 \iff v(x_p) = 0$ for almost all p. In other words, $\langle x \rangle = \frac{x_\infty}{|x_\infty|} \prod_p p^{v(x_p)}$ is a rational number for $x \in \mathbb{A}^\times$, and $x\langle x \rangle^{-1} \in \mathbb{R}_+^\times \times \widehat{\mathbb{Z}}^\times$. □

By translation, we extend the topology on $\widehat{\mathbb{Z}}$ to $\mathbb{A}^{(\infty)} = \mathbb{Q} + \widehat{\mathbb{Z}}$. Therefore $\mathbb{A}$ and $\mathbb{A}^{(\infty)}$ are locally compact rings (and $\mathbb{A}^\times$ is a locally compact group; see [MFG] 3.1, [LFE] 8.1-3 and [EPE] Chapter II for more about the adele topology on the adele points $G(\mathbb{A})$ of an algebraic group G over $\mathbb{Q}$).

Corollary 2.11 *We have* $\mathbb{A}^\times/\mathbb{Q}^\times\mathbb{R}_+^\times \cong \widehat{\mathbb{Z}}^\times$. *This isomorphism* ι *can be normalized so that it takes* $p \in \mathbb{Q}_p \subset \mathbb{A}$ *to* $p^{(p\infty)} = (p, p, \ldots, p, \overset{p}{1}, p, \ldots, p) \in \widehat{\mathbb{Z}}^\times$.

Proof. We see that $\xi \in \mathbb{Q}^\times \cap (\widehat{\mathbb{Z}}^\times \times \mathbb{R}_+^\times)$ means that $\xi = \xi_\infty > 0$ and the numerator and the denominator of $\xi = \xi_p$ is prime to p for all p; so, $\xi = 1$. Thus we conclude $\mathbb{Q}^\times \cap (\widehat{\mathbb{Z}}^\times \times \mathbb{R}_+^\times) = \{1\}$. By the isomorphism theorem:

$$\phi : \mathbb{A}^\times/\mathbb{Q}^\times\mathbb{R}_+^\times = \widehat{\mathbb{Z}}^\times\mathbb{Q}^\times\mathbb{R}_+^\times/\mathbb{Q}^\times\mathbb{R}_+^\times \cong \widehat{\mathbb{Z}}^\times/\mathbb{Q}^\times \cap (\widehat{\mathbb{Z}}^\times \times \mathbb{R}_+^\times) = \widehat{\mathbb{Z}}^\times.$$

Since $p = p^{(p\infty)} p_p p_\infty = 1$ in $\mathbb{A}^\times/\mathbb{Q}^\times\mathbb{R}_+^\times$ for $p_p = (1, \ldots, 1, \overset{p}{p}, 1, \ldots, 1) \in \mathbb{A}^\times$ and $p_\infty = (\overset{\infty}{p}, 1, \ldots, 1) \in \mathbb{A}^\times$, the above isomorphism brings p_p in $\mathbb{Q}_p^\times \subset \mathbb{A}^\times$ to $(p^{(p\infty)})^{-1}$. Defining $\iota(x) = \phi(x)^{-1}$, we get the desired isomorphism. □

Combining this with Theorem 2.8, we get

Theorem 2.12 (Class Field Theory) *We have a canonical (reciprocity) isomorphism* $\iota : GL_1(\mathbb{A}^{(\infty)})/\mathbb{Q}_+^\times = \mathbb{A}^\times/\mathbb{Q}^\times\mathbb{R}_+^\times \cong \mathrm{Gal}(\mathbb{Q}^{cyc}/\mathbb{Q})$ *such that* $p_p \in \mathbb{Q}_p$ *is sent to a Frobenius element at* p, *and the decomposition group at* p *is given by the closure of the image of* $\mathbb{Q}_p^\times$. *Inside the decomposition group, the inertia group is given by the isomorphic image of* $\mathbb{Z}_p^\times$.

Class field theory says slightly more, namely that $\mathbb{Q}^{cyc}$ is the maximal abelian extension $\mathbb{Q}^{ab}$ of $\mathbb{Q}$. For more on this, see [FAN] Chapter 6. Often the above result is expressed as an exact sequence,

$$1 \to \mathbb{Q}^\times\mathbb{R}_+^\times \to \mathbb{A}^\times \xrightarrow{\iota} \mathrm{Gal}(\mathbb{Q}^{ab}/\mathbb{Q}) \to 1,$$

and the map $\iota(x)$ is called the Artin reciprocity map and sometimes written as $\iota(x) = [x, \mathbb{Q}]$. The reciprocity map of the class field theory induces a locally compact (actually compact profinite) topology on the Galois group. This topology can be defined in a way intrinsic to Galois' theory, as we show in the following section.

Exercises

1. Prove that $\Phi_{p^n}(X)$ is irreducible in $\mathbb{Q}[X]$.
2. Give a detailed proof of Corollary 2.4.
3. Prove $\mathbb{Z}_{(q)}[\zeta_{p^n}]$ is integrally closed.
4. Prove that the integer ring O of $\mathbb{Q}[\zeta_N]$ is generated by ζ_N over $\mathbb{Z}$. Hint: First show that $\mathbb{Q}[\zeta_N] = \mathbb{Q}[\zeta_{p^m}] \otimes_{\mathbb{Q}} \mathbb{Q}[\zeta_{q^n}]$ if $N = p^m q^n$ for distinct primes p and q. Then show that $O = \mathbb{Z}[\zeta_{p^m}] \otimes_{\mathbb{Z}} \mathbb{Z}[\zeta_{q^n}]$ using the fact that no prime ramifies in the two fields $\mathbb{Q}[\zeta_{p^m}]$ and $\mathbb{Q}[\zeta_{q^n}]$ at the same time.
5. Give a detailed proof of Theorem 2.8.
6. Deduce the quadratic reciprocity law in the introduction either from Theorem 2.6 or Theorem 2.12.

2.3 A Generalization of Galois Theory

For our later use, we gather here some results from a generalization of Galois' theory given in [IAT] 6.3.

2.3.1 Infinite Galois Extensions

We study the Galois theory when $\dim_K L$ is infinite for a Galois extension L/K. First we recall Galois' fundamental theorem when L/K is finite.

Theorem 2.13 (E. Galois) *Suppose that L/K is a finite Galois extension. Taking an intermediate field $L/M/K$, L/M is a Galois extension and*

(1) *There is one-to-one onto correspondence*

$$\{M : \textit{intermediate field} | L/M/K\} \leftrightarrow \{H : \textit{subgroup of } \mathrm{Gal}(L/K)\}$$

given by $M \mapsto \mathrm{Gal}(L/M)$ and $H \mapsto L^H = \{x \in L | x = \sigma(x) \forall \sigma \in H\}$.

(2) *For two intermediate fields M, M', we have*
- $M \supset M' \iff \mathrm{Gal}(L/M) \subset \mathrm{Gal}(L/M')$;
- $\mathrm{Gal}(M \cap M')$ *is the subgroup of* $\mathrm{Gal}(L/K)$ *generated by the two subgroups* $\mathrm{Gal}(L/M)$ *and* $\mathrm{Gal}(L/M')$;
- $\mathrm{Gal}(L/MM') = \mathrm{Gal}(L/M) \cap \mathrm{Gal}(L/M')$;
- *For* $\sigma \in \mathrm{Gal}(L/K)$, $\mathrm{Gal}(L/\sigma(M)) = \sigma \mathrm{Gal}(L/M)\sigma^{-1}$.

(3) *M/K is a Galois extension if and only if* $\mathrm{Gal}(L/M)$ *is a normal subgroup in the Galois group* $\mathrm{Gal}(L/K)$. *In particular, by $\sigma \mapsto \sigma|_M$, we have an isomorphism:* $\mathrm{Gal}(L/K)/\mathrm{Gal}(L/M) \cong \mathrm{Gal}(M/K)$.

We admit this theorem, just referring the reader to any book on elementary Galois theory.

What we want to do in this subsection is to generalize the above theorem to an infinite Galois extension. Hereafter we suppose that L/K is an infinite Galois extension. Since each element $\xi \in L$ satisfies a polynomial equation of

finite degree with coefficients in K, there are only finitely many conjugates of ξ. Thus the Galois closure $K[\xi]^{gal}$ of $K[\xi]$ over K is a finite Galois extension of K, and we get from Theorem 2.13 (2),

$$L = \bigcup_{\xi \in L} K[\xi]^{gal} \Leftrightarrow \bigcap_{\xi \in L} \mathrm{Gal}(L/K[\xi]^{gal}) = \{\mathrm{id}\}. \tag{2.4}$$

We first generalize the third assertion of the above theorem.

Lemma 2.14 *Suppose M is an intermediate Galois extension of a Galois extension L/K. Then the group $\mathrm{Gal}(L/M)$ is a normal subgroup of $\mathrm{Gal}(L/K)$, and $\sigma \mapsto \sigma|_M$ induces an isomorphism $\mathrm{Gal}(L/K)/\mathrm{Gal}(L/M) \cong \mathrm{Gal}(M/K)$. In particular, if M/K is a finite extension, then we have the equality $[\mathrm{Gal}(L/K) : \mathrm{Gal}(L/M)] = [M : K] < \infty$.*

Proof. If M/K is a finite extension, for any finite Galois extension L' with $L \supset L' \supset M$, we can extend $\sigma \in \mathrm{Gal}(M/K)$ to $\tau \in \mathrm{Gal}(L'/K)$ (assertion (3) of Theorem 2.13). Since L is a union of finite Galois extensions of L', $\sigma \in \mathrm{Gal}(L'/K)$ extends to $\tau \in \mathrm{Gal}(L/K)$. If M/K is an infinite extension, writing M as a union of finite Galois extensions M'/K and applying the above fact to M'/K, we find that any $\sigma \in \mathrm{Gal}(M/K)$ can be extended to $\tau \in \mathrm{Gal}(L/K)$. Thus $\sigma \mapsto \sigma|_M$ is a surjective homomorphism of $\mathrm{Gal}(L/K)$ onto $\mathrm{Gal}(M/K)$. Its kernel is given by $\mathrm{Gal}(L/M)$ which is normal. □

Basically the same proof as above yields the following result which looks stronger than the lemma.

Corollary 2.15 *Let M be an intermediate extension of L/K. Then $\sigma \mapsto \sigma|_M$ induces an isomorphism $\mathrm{Gal}(L/K)/\mathrm{Gal}(L/M) \cong \mathrm{Hom}_K(M, L)$, where the set $\mathrm{Hom}_K(M, L)$ is the collection of K-linear field homomorphisms of M into L.*

The proof of this corollary is left to the reader as an exercise (Exercise 1).

We now give a topology on the group $\mathrm{Gal}(L/K)$ in the following way. Define a fundamental system of open neighborhoods of the identity by the collection of subgroups of the form $\mathrm{Gal}(L/M)$ for a finite extension M/K. If $H = \mathrm{Gal}(L/M)$ for a finite extension M/K, the Galois closure M^{gal} is still a finite Galois extension of K; so, we find a normal open subgroup $\mathrm{Gal}(L/M^{gal})$ inside any open subgroup H. Thus we may define the system to be the set $\mathfrak{U}$ of all normal subgroups $N = \mathrm{Gal}(L/M)$ for finite Galois extensions M/K inside L. For each $\sigma \in \mathrm{Gal}(L/K)$, we define the system of neighborhoods of σ to be $\sigma\mathfrak{U} = \mathfrak{U}\sigma = \{\sigma N = N\sigma | N \in \mathfrak{U}\}$. By this, if $\mathfrak{U}$ satisfies the axiom of a fundamental system of open neighborhoods of a point, G becomes a topological group, that is, G is a group with a topology for which the multiplication $(a, b) \mapsto ab$ and inverse $a \mapsto a^{-1}$ are continuous. The axiom is checked by

Lemma 2.16 *We have*

(1) *For $\sigma \neq 1$, there exists $N \in \mathfrak{U}$ with $\sigma \notin N$ (so, G is a Hausdorff group);*

(2) $N, H \in \mathfrak{U} \Rightarrow N \cap H \in \mathfrak{U}$.

Proof. By (2.4), $L = \bigcup_M M$ for finite Galois extensions M/K. If $\sigma \in \bigcap_M \mathrm{Gal}(L/M)$, σ is the identity over all M; so, $\sigma = 1$ in $\mathrm{Gal}(L/K)$. Thus there exists a finite Galois extension M/K inside L such that $\sigma \notin N = \mathrm{Gal}(L/M)$ if $\sigma \neq 1$. By definition, we have $N \in \mathfrak{U}$; so, assertion (1) follows.

Write $N = \mathrm{Gal}(L/M)$ and $H = \mathrm{Gal}(L/M')$. Then MM' is a composite of two finite Galois extensions; so, it is a finite Galois extension. By definition, σ is trivial over $MM' \iff \sigma$ is trivial over M and M'. Then by Lemma 2.14, we have $\mathrm{Gal}(L/MM') = N \cap H$, and hence $N \cap H \in \mathfrak{U}$, which proves (2). □

The topology we have defined was first introduced by Krull, and therefore it is called the Krull topology on $\mathrm{Gal}(L/K)$.

Proposition 2.17 *For an infinite Galois extension L/K, $\mathrm{Gal}(L/K)$ is a compact group under the Krull topology.*

Proof. Since $\mathrm{Gal}(L/K)$ is Hausdorff, we need to show that for any given infinite subset Σ in $\mathrm{Gal}(L/K)$, there exists an element $\sigma \in \mathrm{Gal}(L/K)$ such that $\Sigma \cap H\sigma$ is infinite for all $H \in \mathfrak{U}$; that is, Σ has a limit point $\sigma \in \mathrm{Gal}(L/K)$. Since $H = \mathrm{Gal}(L/M)$ for a finite Galois extension M/K, there are finitely many automorphisms of M/K. Thus there exists an infinite subset Σ_H of Σ that induces a single automorphism σ_M on M. For any extension $\sigma' \in \Sigma_H$ of σ_M to L, the coset $H\sigma'$ is the collection of $\tau \in \mathrm{Gal}(L/K)$ (by Lemma 2.14) with $\tau|_M = \sigma_M$. We make the construction of σ_M compatible with an increasing sequence of Galois extensions M_j of K inside L. In other words, we find a sequence of finite Galois extensions M_j/K $(j = 1, 2, \dots)$ with $K \subset M_j \subset M_{j+1} \subset L$ and $\sigma_j = \sigma_{M_j} \in \mathrm{Gal}(M_j/K)$ such that $L = \bigcup_j M_j$ and $\sigma_{j+1}|_{M_j} = \sigma_j$ for all j. We have found an element $\sigma \in \mathrm{Gal}(L/K)$ such that $\sigma|_{M_j} = \sigma_{M_j}$ for all j. This σ is a limit point of Σ. □

Corollary 2.18 *If M is an intermediate field of an infinite Galois extension L/K, $\mathrm{Gal}(L/M)$ is a closed subgroup of $\mathrm{Gal}(L/K)$.*

Proof. For any finite Galois extension M'/K inside L, we have by definition the following identity: $i^{-1}(\mathrm{Gal}(L/MM')) = \mathrm{Gal}(L/MM')$ for the inclusion map $i : \mathrm{Gal}(L/M) \hookrightarrow \mathrm{Gal}(L/K)$. For any $N \in \mathfrak{U}$, we find a finite Galois extension M'/K inside L such that $N = \mathrm{Gal}(L/M')$. Since MM'/M is a finite Galois extension, $i^{-1}(N) = \mathrm{Gal}(L/MM')$ is open in $\mathrm{Gal}(L/M)$; so, i is continuous. The image of a compact set by a continuous map is compact; in particular, it is closed. □

We now prove a generalization of Theorem 2.13.

Theorem 2.19 *Let L/K be an infinite Galois extension. We have the following canonical one-to-one onto correspondence,*

$$\{\textit{intermediate fields } M \textit{ of } L/K\} \leftrightarrow \{\textit{closed subgroups } H \textit{ of } \mathrm{Gal}(L/K)\}$$

induced by $M \mapsto \mathrm{Gal}(L/M)$ and $H \mapsto L^H$. Moreover,

(1) *Every open subgroup of* $\mathrm{Gal}(L/K)$ *is closed, and open subgroups correspond to finite extensions* L^H/K.
(2) *Let* $\{N_\iota\}_{\iota\in I}$ *for an index set* I *be a collection of closed subgroups of the Galois group* $\mathrm{Gal}(L/K)$. *Writing* $N_\iota = \mathrm{Gal}(L/K_\iota)$, *for the composite* M *of all* $\{K_\iota\}_{\iota\in I}$, *we have* $\bigcap_{\iota\in I} N_\iota = \mathrm{Gal}(L/M)$.
(3) *If* $\{K_\iota\}_{\iota\in I}$ *is a collection of intermediate fields of* L/K, *then the group* $\mathrm{Gal}(L/\cap_{\iota\in I} K_\iota)$ *is the closure of the subgroup in* $\mathrm{Gal}(L/K)$ *generated by* $\mathrm{Gal}(L/K_\iota)$ *for all* i.
(4) *For* $\sigma \in \mathrm{Gal}(L/K)$, *we have* $\mathrm{Gal}(L/\sigma(M)) = \sigma \cdot \mathrm{Gal}(L/M)\sigma^{-1}$.
(5) *If* M/K *is a Galois extension inside* L, *then* $\mathrm{Gal}(L/M)$ *is a normal closed subgroup and* $\mathrm{Gal}(L/K)/\mathrm{Gal}(L/M) \cong \mathrm{Gal}(M/K)$ *induced by* $\sigma \mapsto \sigma|_M$ *is an isomorphism of topological groups.*

Proof. Let $\mathfrak{M}$ be the set of all intermediate fields of L/K and $\mathfrak{H}$ be the set of all closed subgroups of $\mathrm{Gal}(L/K)$. For any $M \in \mathfrak{M}$, $\mathrm{Gal}(L/M)$ is a closed subgroup of $\mathrm{Gal}(L/K)$ by Corollary 2.18; so, $M \mapsto \mathrm{Gal}(L/M)$ defines a map $\mathfrak{M} \to \mathfrak{H}$. We define a reverse map $\mathfrak{H} \to \mathfrak{M}$ by $H \mapsto L^H$.

We first prove that $\mathrm{Gal}(L/L^H) = H$. Since $\mathrm{Gal}(L/L^H) \subset \mathrm{Gal}(L/K)$ is the collection of all automorphisms of L fixing L^H, we confirm that $H \subset \mathrm{Gal}(L/L^H)$. Conversely, for each finite Galois extension M/L^H inside L, H acts on M nontrivially if $M \neq L^H$, and the image under $\sigma \mapsto \sigma|_M$ in $\mathrm{Gal}(M/L^H)$ is a subgroup H' of $\mathrm{Gal}(M/L^H)$ and $L^H = M^{H'}$. By Theorem 2.13 (1), $H' = \mathrm{Gal}(M/L^H)$. In other words, for each $\sigma \in \mathrm{Gal}(L/L^H)$, we find $h_M \in H$ such that $h_M|_M = \sigma|_M$. We consider the infinite set $\Sigma = \{h_M\}_M \subset H$ with M running through all finite Galois extensions M/L^H. Since H is a closed subgroup of a compact group $\mathrm{Gal}(L/K)$, H itself is compact. Then Σ has a unique accumulation point h. Then for each finite Galois extension M/L^H, $\sigma|_M = h|_M$. Since $L = \bigcup_M M$ for finite Galois extensions of M/L^H, we find that $h = \sigma$ and $H = \mathrm{Gal}(L/L^H)$.

We now prove $L^{\mathrm{Gal}(L/M)} = M$. By Corollary 2.18, $H = \mathrm{Gal}(L/M) \in \mathfrak{H}$. By definition, $M \subset L^H$. Supposing that $L^H \neq M$, we try to get a contradiction. Pick $\xi \in L^H - M$. Then $M[\xi]/M$ is a finite extension. Thus $M[\xi]^{gal} \subset L$ because L/K is normal. Since $\xi \notin M$, $\mathrm{Gal}(M[\xi]^{gal}/M) \neq \{1\}$. Pick $\sigma \in \mathrm{Gal}(M[\xi]^{gal}/M)$ with $\sigma(\xi) \neq \xi$. Then by Lemma 2.14, we find $\tau \in \mathrm{Gal}(L/M) = H$ such that $\tau(\xi) = \sigma(\xi) \neq \xi$. This is wrong since $\sigma(\xi)$ has to be ξ because $\xi \in L^H$. Thus we find that $L^{\mathrm{Gal}(L/M)} = M$, and the correspondence is one-to-one and onto.

For an open subgroup H, $\mathrm{Gal}(L/K)/H$ is discrete and compact; so, it is finite (see Exercise 4). Since H is the kernel of the continuous map $\mathrm{Gal}(L/K) \twoheadrightarrow \mathrm{Gal}(L/K)/H$, H is closed. By $H = \mathrm{Gal}(L/L^H)$, we see from Lemma 2.14 and Corollary 2.15 that $\mathrm{Gal}(L/K)/H = \mathrm{Hom}_K(L^H, L)$, and hence L^H/K is a finite extension by Corollary 2.15. This shows (1).

Assertion (2) follows from the fact: $\sigma \in \mathrm{Gal}(L/K)$ is the identity on every K_ι if and only if σ is the identity on the composite of all K_i (Exercise 2).

Let $\mathrm{Gal}(L/\cap_{\iota\in I} K_\iota) = H'$. Then by definition, H' is a closed subgroup containing all $H_\iota = \mathrm{Gal}(L/K_\iota)$, and it contains the closure H of the subgroup generated by H_ι for all i. Since H is a closed subgroup (so in $\mathfrak{H}$) containing H_ι, we see $L^H \subset K_\iota$ and hence $L^H \subset \bigcap_{\iota\in I} K_\iota$. Thus $H \supset H' = \mathrm{Gal}(L/\cap_{\iota\in I} K_i)$. This shows $H = H'$.

Assertion (4) is plain from the definition of the Galois group, and the last assertion follows from Lemma 2.14. □

2.3.2 Automorphism Group of a Field

We study $\mathrm{Aut}(L/K)$ for an arbitrary extension L/K of fields. An intermediate extension M of L/K is called *of finite type* over K if there are finitely many elements $x_1, \ldots, x_n$ that generate M as a field over K. We define a fundamental system $\mathfrak{U}$ of open neighborhoods of 1 on $\mathrm{Aut}(L/K)$ by

$$\mathfrak{U} = \{H = \mathrm{Aut}(L/M) | M/K \text{ is of finite type}\}. \tag{2.5}$$

Since a composite of two intermediate fields of finite type is again of finite type, $\mathfrak{U}$ satisfies the axiom of the fundamental system of neighborhoods (see the proof of Lemma 2.16 and Exercise 3). Then define a system of neighborhoods of $\sigma \in \mathrm{Aut}(L/K)$ by $\sigma\mathfrak{U} \cup \mathfrak{U}\sigma$. This topology gives a topological group structure on $\mathrm{Aut}(L/K)$. Let $\mathfrak{H}$ be the set of all compact subgroups of $\mathrm{Aut}(L/K)$, and let $\mathfrak{M}$ be the set of all intermediate fields M in L/K such that L/M is a Galois extension. Then we have

Proposition 2.20 (Jacobson) *We have a canonical bijective correspondence $\mathfrak{M} \cong \mathfrak{H}$ given by $M \mapsto \mathrm{Gal}(L/M)$ and $H \mapsto L^H$.*

Proof. Pick a compact subgroup H of $\mathrm{Aut}(L/K)$. Then for each $\xi \in L$, the topological group $\mathrm{Aut}(L/K(\xi))$ is an open subgroup of $\mathrm{Aut}(L/K)$. Thus $H' = H \cap \mathrm{Gal}(L/K(\xi))$ is an open subgroup of H. In particular, H/H' is compact and discrete; so, it is finite (Exercise 4). Since $f(X) = \prod_{h\in H/H'}(X - h(\xi))$ has coefficients in L^H, ξ is algebraic over L^H. Since $h(\xi) = h'(\xi) \Leftrightarrow hH' = h'H'$, we find that the roots of $f(X)$ are all distinct; so $L^H(\xi)/L^H$ is a separable extension. Since $h(\xi) \in L$ for all $h \in H$, all the roots of $f(X)$ are in L. The Galois closure M of $L^H(\xi)$ over L^H is a Galois extension of L^H inside L. Since $L = \bigcup_{\xi\in L} L^H(\xi)$, we find that L/L^H is a Galois extension. Then by Theorem 2.19, we find that $H = \mathrm{Gal}(L/L^H)$. Starting from $M \in \mathfrak{M}$, we find that $\mathrm{Gal}(L/M) \in \mathfrak{H}$ and $M = L^{\mathrm{Gal}(L/M)}$ by Theorem 2.19. □

A field extension L/K is of finite type if L is generated over K by finitely many elements (as a field). Let $\mathfrak{H}^o$ be the set of all open compact subgroups of $\mathrm{Aut}(L/K)$ and $\mathfrak{M}^o$ be the subset of $\mathfrak{M}$ made up of fields M that are of finite type over K.

Corollary 2.21 *If $\mathfrak{M}^o$ is nonempty, then $\mathrm{Aut}(L/K)$ is locally compact, and the one-to-one onto correspondence in* Proposition 2.20 *induces $\mathfrak{M}^o \cong \mathfrak{H}^o$.*

Proof. If $\mathfrak{M}^o$ is nonempty, pick $M \in \mathfrak{M}^o$. Then the identity of $\mathrm{Aut}(L/K)$ has an open neighborhood $\mathrm{Gal}(L/M)$, which is also compact by Proposition 2.17. By translating this neighborhood to any point σ on $\mathrm{Aut}(L/K)$ through multiplication by σ, every point has an open-compact neighborhood; so, $\mathrm{Aut}(L/K)$ is locally compact. We need to show that L^H/K is of finite type over K if $H \in \mathfrak{H}^o$. Consider $F = ML^H$. Then L/F is a Galois extension; so, $H' = \mathrm{Gal}(L/F)$ is a compact subgroup of $\mathrm{Gal}(L/M)$. Since F/M is a finite extension, F is of finite type over M and hence over K. Thus the subfield L^H of a field F is of finite type over K. □

Exercises

1. Prove Corollary 2.15.
2. Give a detailed proof of the assertions (2) and (4) of Theorem 2.19.
3. Prove that the Krull topology on $\mathrm{Aut}(L/K)$ given by (2.5) is well-defined.
4. Prove that a compact and discrete set is a finite set.

2.4 Algebraic Curves over a Field

Here we give an exposition of classical theory of algebraic curves over a field from the viewpoint of the theory of discrete valuation. See [ALF] for a more detailed exposition on algebraic curves from this point of view. Under this setting, in the following section, we give a prototype of classification problems of abelian varieties (description of the moduli of elliptic curves over a field). To treat classification problems over rings, a more sophisticated language, for example, the language of schemes, is necessary. From the next chapter on, we treat the general case via the theory of schemes.

2.4.1 Algebraic Function Fields

Let K be a field of an arbitrary characteristic. An algebraic function field $\mathfrak{K}$ of dimension 1 is a finitely generated nonalgebraic field extension $\mathfrak{K}/K$ such that for any $x \in \mathfrak{K}$ transcendental over K, $\mathfrak{K}/K(x)$ is algebraic. We assume that the algebraic closure of K in $\mathfrak{K}$ is K itself. In this case, we say that $\mathfrak{K}$ is defined over K. Replacing K by its algebraic closure in $\mathfrak{K}$, we may always assume that $\mathfrak{K}$ is defined over K.

We relate field theory to the geometric theory of projective algebraic curves by considering local coordinates. A prototypical example of algebraic function fields is given by the meromorphic function field over $\mathbb{C}$ of a compact Riemann surface. Each point of a Riemann surface has a coordinate neighborhood, and the collection of all coordinate neighborhoods by definition determines the Riemann surface. Each meromorphic function has Laurent expansion at a given point, which gives a well-defined order of vanishing (or order of pole) of

the function at the given point. Associating the order at a point with meromorphic functions gives a valuation of the function field specific to the point. An idea of how to algebraize a Riemann surface is to consider the set of all valuations of its function field trivial over the base field and to replace coordinate neighborhoods by corresponding valuations. We recall here a formal definition of discrete valuation rings (DVR) inside an algebraic function field $\mathfrak{K}$ over K. A *discrete valuation* v of $\mathfrak{K}$ trivial on K is a surjective map $v : \mathfrak{K} \to \mathbb{Z} \sqcup \{\infty\}$ satisfying the following four conditions:

(V0) $v(K^\times) = 0$ (triviality over the base field);
(V1) $v(f) = \infty \Leftrightarrow f = 0$;
(V2) $v(f+g) \geq \min(v(f), v(g))$ for all $f, g \in \mathfrak{K}$;
(V3) $v(fg) = v(f) + v(g)$ for all $f, g \in \mathfrak{K}$.

Here we agree as a convention to have $a + \infty = \infty$ and $\infty > a$ for all $a \in \mathbb{Z}$. By the above properties, $\mathcal{V}_v = \{f \in \mathfrak{K} | v(f) \geq 0\}$ is a subring of $\mathfrak{K}$, and either $x \in \mathfrak{K}$ or $\frac{1}{x}$ belongs to $\mathcal{V}_v$. In particular, the field of fractions of $\mathcal{V}_v$ is equal to $\mathfrak{K}$. The ring $\mathcal{V}_v$ is called a *discrete valuation ring*, and $\mathfrak{m} = \{f \in \mathcal{V}_v | v(f) \geq 1\}$ is a unique maximal ideal of $\mathcal{V}_v$ (so, $\mathcal{V}_v$ is local). The field $\mathcal{V}_v/\mathfrak{m}$ is called the residue field of the valuation v and is a finite extension of the base field K. Moreover, every ideal of $\mathcal{V}_v$ (except for the zero ideal) is a power of $\mathfrak{m}$, and $\mathfrak{m}^n = \{f \in \mathfrak{K} | v(f) \geq n\}$. See [CRT] Chapter 4 for more about valuation rings.

Our goal in this subsection is to create a space (a geometric object) from the purely algebraic notion of algebraic function fields $\mathfrak{K}$. The object is called the Zariski–Riemann space and is the collection of all valuations of $\mathfrak{K}$ trivial over the base field K. The space is an algebraic replacement of the associated Riemann surface for $K = \mathbb{C}$.

Example 2.22 We start with the simplest Riemann surface: the Riemann sphere $\mathbf{P} = \mathbf{P}^1 = \mathbb{C} \cup \{\infty\}$. The meromorphic function field of the sphere $\mathbf{P}$ is isomorphic to the rational function field $\mathfrak{K} = \mathbb{C}(x)$. Then the polynomial ring $\mathbb{C}[x]$ corresponds to the Riemann sphere $\mathbf{P}$ with coordinate x in the following sense. The space $\mathbf{P}$ is covered by two coordinate neighborhoods U_0 and U_∞ identical to $\mathbb{C}$, U_0 is centered at 0 with complex coordinate x, and the other U_∞ is centered at ∞ with coordinate $x' = \frac{1}{x}$. Then the field of meromorphic functions of $\mathbf{P}$ is given by $\mathbb{C}(x)$, and the polynomial ring $\mathbb{C}[x]$ is the ring of functions with the only possible pole at ∞. Each meromorphic function ϕ on U_0 finite at 0 has its Taylor expansion in the coordinate x: $\phi(x) = \sum_{n=0}^{\infty} a_n x^n$ whose radius of convergence is positive. If ϕ is not finite at 0, for some positive m, $x^m \phi(x)$ is finite at 0, and $\phi(x)$ therefore has its Laurent expansion $\sum_{n \geq -m} a_n x^n$. We can define a valuation v_0 on $\mathbb{C}(x)$ by assigning the exponent of the leading term of the Laurent expansion to a given meromorphic function $\phi \in \mathbb{C}(x)$. The valuation $v_0(\phi)$ is just the zero order of the function ϕ holomorphic at 0. For any other point $\alpha \in \mathbb{C} = U_0$ (resp. $\alpha = \infty \in U_\infty$), we can take $t_\alpha = x - \alpha$ (resp. $t_\infty = x^{-1}$) as a coordinate around α, and we can think of the valuation v_α giving the order of zero at α. In differential geometry or in complex analysis, local behavior of functions reflected by Taylor

expansion often determines in the aggregate global properties of the function. Partial fraction expansion determines elements in $\mathbb{C}(x)$; so, local knowledge of the valuation $v_\alpha(\phi)$ almost determines $\mathbb{C}(x)$. Thus it is natural to expect that the set Z of valuations $\{v_\alpha\}_\alpha$ determine $\mathbf{P}$. Indeed, at least set-theoretically, Z is in bijection with $\mathbf{P}$.

We choose a transcendental element $x \in \mathfrak{K}$ so that $\mathfrak{K}$ is a finitely generated separable algebraic extension of $K(x)$. Then $\mathfrak{K}/K(x)$ is a simple extension; in other words, we can find a single generator $y \in \mathfrak{K}$ such that $\mathfrak{K} = K(x)[y]$. We may assume that y is integral over $K[x]$ (because some multiple of y by an element in $K[x]$ is integral over $K[x]$), and y satisfies an equation $y^m + a_1(x)y^{m-1} + \cdots + a_m(x) = 0$.

Suppose that $K = \mathbb{C}$. Take a general compact Riemann surface R. Then on the given Riemann surface R, choosing two generators (x, y) in its meromorphic function field, the relation of (x, y) gives a projective algebraic curve V in $\mathbf{P}^2$. Moreover, defining the algebraic function field $\mathbb{C}(V)$ of V by the collection of nontrivial restriction to V of rational functions $P(x, y)/Q(x, y)$ (with two polynomials P and Q of equal degree), the meromorphic function field of R is isomorphic to the algebraic function field $\mathbb{C}(V)$ of V.

Here is another slightly more nontrivial example.

Example 2.23 Consider the function field $\mathfrak{K} = \mathbb{C}(x)[y]$ defined by the equation $y^2 = x(x-1)(x-\lambda)$ for $\lambda \in \mathbb{C}$ different from 0 and 1. We consider the square root $y(x) = \sqrt{x(x-1)(x-\lambda)}$ on U_0. This function has two values on $\mathbf{P}$ vanishing at $0, 1, \lambda$ and has a pole at ∞; in other words, its inverse vanishes at ∞. Thus the locus of the point $(x, y, 1) \in \mathbf{P}^2$ is a two-sheet covering of $\mathbf{P}$ that ramifies at $0, 1, \lambda, \infty$. Around $x = 0$, if a point circles around $(x, y) = (0, 0)$, for two values of $\pm y$, we have one value of x. In other words, cutting a line segment $[0, 1]$ and $[\lambda, \infty]$ from two copies of $\mathbf{P}$ and gluing the two corresponding segments, we get the donut-shape Riemann surface R, on which the functions y and x both have single values. The field of meromorphic functions over R is given by $\mathbb{C}(x)[y]$. Moreover, we can embed R into $\mathbf{P}^2$ by $P \mapsto (x(P), y(P))$, which satisfies the given equation $y^2 = x(x-1)(x-\lambda)$. In other words,

$$R \cong V = \left\{(x, y) \in \mathbf{P}^2 \middle| y^2 = x(x-1)(x-\lambda)\right\},$$

because x moves around all possible values in $\mathbf{P} = \mathbb{C} \cup \{\infty\}$.

For each point $\alpha \in \mathbb{C}$ different from $0, 1, \lambda$, pick a point P from the two points P, Q of R over α. Then P has an open neighborhood U with coordinate $t = x - \alpha$. Any meromorphic function f defined over U has a Laurent expansion $f(t) = \sum_{n \gg -\infty} a_n t^n$. We define a valuation $v_P : \mathbb{C}(V) \to \mathbb{Z} \cup \{\infty\}$ given by

$$v_P(f) = m \quad \text{if } a_m \neq 0 \text{ and } a_n = 0 \text{ if } n < m.$$

Thus we have (V0–3) for v_P and

$$\mathcal{V}_P = \left\{ f \in \mathfrak{K} \middle| v_P(x) \geq 0 \right\} \tag{2.6}$$

is a valuation ring (Exercise 1) with $v_P(\mathbb{C}^\times) = 0$ (i.e., v_P is trivial on $\mathbb{C}$).

For P over one of the four ramified points $\alpha = 0, 1, \lambda, \infty$, for an open neighborhood O of α that does not contain any of $0, 1, \lambda, \infty$ different from α, $x^{-1}(O) = U$ is an open neighborhood of P and y gives a local coordinate t if $\alpha \neq \infty$ and y^{-1} gives a local coordinate t of U if $\alpha = \infty$. Then we define $\mathcal{V}_P$ exactly in the same way. By definition, for each $x \in \mathfrak{K}$, either $v_P(x) \geq 0$ or $v_P(x) < 0$; in other words, x or x^{-1} is in $\mathcal{V}_P$. Thus we have $\mathfrak{K} = Frac(\mathcal{V}_P) = \{ \frac{a}{b} | a \in \mathcal{V}_P, b \in \mathcal{V}_P - \{0\} \}$.

Start conversely with a discrete valuation $v : \mathfrak{K} \to \mathbb{Z} \cup \{\infty\}$ with $v(\mathbb{C}^\times) = 0$. Then $\mathcal{V} = \{ f \in \mathfrak{K} | v(f) \geq 0 \}$ is a discrete valuation ring. Again for each $x \in \mathfrak{K}$, either $x \in \mathcal{V}$ or $x^{-1} \in \mathcal{V}$. First suppose that $x \in \mathcal{V}$. Then $\mathbb{C}[x] \subset \mathcal{V}$. Let $\mathfrak{m} = \{ f \in \mathcal{V} | v(f) > 0 \}$. Then $\mathfrak{m}$ is the maximal ideal of $\mathcal{V}$, and $\mathcal{V}/\mathfrak{m}$ is a field extension of $\mathbb{C}$. Since $\mathbb{C}[x]$ is a principal ideal domain, $\mathfrak{m} \cap \mathbb{C}[x]$ is either $(x - \alpha)$ for $\alpha \in \mathbb{C}$ or (0). If $\mathfrak{m} \cap \mathbb{C}[x] = (0)$, we find that $\mathcal{V}/\mathfrak{m}$ contains an isomorphic image of $\mathbb{C}[x]$. Thus the transcendental degree of $\mathcal{V}/\mathfrak{m}$, (i.e., $\dim(\mathcal{V}/\mathfrak{m})$) is larger than or equal to 1. Since $\mathcal{V}$ has Krull dimension 1 (cf. [CRT] Section 5), this implies $\mathfrak{m} = (0)$, which is impossible. We find that $\mathfrak{m} \cap \mathbb{C}[x] = (x - \alpha)$. This implies that $\mathcal{V}/\mathfrak{m} = \mathbb{C}$ and $y \mod \mathfrak{m} = \beta$ which satisfies $\beta^2 = \alpha(\alpha - 1)(\alpha - \lambda)$. In other words, taking the point P with coordinate $(\alpha, \beta) \in V$, we find $\mathcal{V} = \mathcal{V}_P$.

If $x \notin \mathcal{V}$, we find $z' = \frac{1}{x} \in \mathcal{V}$. Then we see easily that $\mathcal{V} = \mathcal{V}_\infty$. We have found the following fact:

$$R \cong \left\{ \mathcal{V} \middle| \mathcal{V} \text{ is a DVR with a valuation trivial on } \mathbb{C} \right\}.$$

This is an algebraic interpretation by Oscar Zariski of a Riemann surface as a space of all valuations trivial on $\mathbb{C}$, and the space at the right-hand side of the above identity is called a *Zariski–Riemann* space. It is intriguing that all rational primes correspond to discrete valuations of $\mathbb{Q}$, and we might want to think that $\mathbb{Q}$ is an arithmetic analogue of a Riemann surface.

For a smooth projective curve V defined over a field K, we write $K(V)$ for its field of K-rational meromorphic functions, and $V(M)$ denotes the set of M-rational points of the curve V for an extension M/K. For a point P of V algebraic over K, we write $K(P)$ for the field generated over K by the coordinates of P.

Theorem 2.24 *For a given algebraic function field $\mathfrak{K}$ defined over a field K, there exists a unique smooth projective curve V defined over K such that $\mathfrak{K} \cong K(V)$. We also have a canonical one-to-one onto correspondence:*

$$V(K) \cong \{ \textit{discrete valuations of } \mathfrak{K} \textit{ trivial over } K^\times \textit{ with residue field } K \}.$$

If $\mathfrak{K}'/\mathfrak{K}$ is an extension of algebraic function fields defined over K, the corresponding smooth projective algebraic curve gives rise to a covering $\pi : V' \to V$

such that $f \mapsto f \circ \pi$ gives the inclusion $\mathfrak{K} \hookrightarrow \mathfrak{K}'$. In particular $\mathcal{V}_P \cap \mathfrak{K} = \mathcal{V}_{\pi(P)}$ for the valuation ring $\mathcal{V}_P$ corresponding to $P \in V'$.

Proof. We generalize the argument (in the two examples) of making a projective algebraic curve out of valuations of $\mathfrak{K}$ for a general algebraic function field over a field K. We start with $K(\mathbf{P}^1) = K(x)$ with one transcendental element x. The main tool is the fact that $K[x]$ is a principal ideal domain. Pick a nonzero prime ideal P of $K[x]$. Pick $f \in P$, and factorize f into a product $f = c\prod_t t(x)^{e(t)}$ of irreducible monic polynomials $t(x)$ with a constant $c \in K^\times$. Since P is a prime, one of the irreducible factors $t(x)$ has to be in P. Then $K(P) = K[x]/P$ is covered by $K[x]/(t(x))$, which is a finite extension of K. We get a surjective field homomorphism $\pi : K[x]/(t(x)) \twoheadrightarrow K(P) = K[x]/P$. Since a field homomorphism is always injective (a field only has two ideals: (0) or itself), we find $K(P) = K[x]/(t(x))$ and $P = (t(x))$. So any nonzero prime ideal is a maximal ideal. We define the valuation $v_P : K(x) \to \mathbb{Z} \cup \{\infty\}$ by $v_P(f) = v_P(c\prod_t t(x)^{e(t)}) = e(t)$. Then the valuation ring $\mathcal{V}_P = \{f \in K(x) | v_P(f) \geq 0\}$ is associated with the point P by $\mathfrak{m}_P \cap K[x] = P$ for the maximal ideal $\mathfrak{m}_P = \{f \in \mathcal{V}_P | v_P(f) > 0\}$.

We study $\mathcal{V}_P/\mathfrak{m}_P^n \cong K[x]/P^n$, because $\mathfrak{m}_P^j/\mathfrak{m}_P^{j+1} = (K[x]/P)t(x)^j = P^j/P^{j+1}$. For a given $f \in K[x]$, by the division algorithm, we find a quotient $q \in K[x]$ and a remainder r such that $f = tq + r$ with $\deg(r) < \deg(t(x))$; so, we may regard r as a unique element in $K(P)$. We again apply the division algorithm to q, getting $q = tq_1 + r_1$; in other words, $f = r + r_1 t + t^2 q_1$. Repeating this process, we can expand f into: $f = r + r_1 t + r_2 t^2 + r_3 t^3 + \cdots$. This appears to be an infinite series, but is actually a finite sum by comparing the degrees of the both sides. We write $\widehat{\mathcal{V}}_P = K(P)[[t]]$, which is a formal power series ring, and embed $\mathcal{V}_P$ into $\widehat{\mathcal{V}}_P$ by the above expansion; so, the image of $\mathcal{V}_P$ is made up of polynomials in t with coefficients in $K(P)$. If one introduces the projective limit of $\mathcal{V}_P/\mathfrak{m}_P^n$, we find $\widehat{\mathcal{V}}_P = \varprojlim_n \mathcal{V}_P/\mathfrak{m}_P^n$.

For a given valuation ring $\mathcal{V} = \{f \in K(x) | v(f) \geq 0\}$ with $v(K^\times) = 0$ (but $v(K(x)) = \mathbb{Z} \cup \{\infty\}$), we find either $x \in \mathcal{V}$ or $x \notin \mathcal{V}$. If $x \in \mathcal{V}$, then $K[x] \subset \mathcal{V}$; so, v is a valuation on $K[x]$. Since v cannot be trivial over $K[x]$ (otherwise, it is trivial on $K(x)$), the intersection $\mathfrak{m} \cap K[x]$ is a prime ideal P of $K[x]$. Then $v = v_P$ because $v(f) \geq m \iff f \in \mathfrak{m}^m \cap K[x] = P^m$ because P is principal. If $x \notin \mathcal{V}$, then $x^{-1} \in \mathcal{V}$, and for the prime $Q = (x^{-1})$ in $K[x^{-1}]$, we have $K[x^{-1}]_Q = \mathcal{V}$, and $\mathcal{V}$ corresponds to the point $\infty = Q$.

If K is algebraically closed, each nonzero prime ideal of $K[x]$ is the form of $(x - \alpha)$ with $\alpha \in K$, which tells us

$$\{\mathcal{V} | \text{valuation rings of } K(x) \text{ trivial on } K\} \cong K \cup \{\infty\} = \mathbf{P}(K) \quad \text{(canonically)}.$$

If K is not algebraically closed, take an algebraic closure $\overline{K}$ of K. Defining $\mathbf{P}(K)$ by the set of points in $\mathbf{P}(\overline{K})$ with coordinates in K, we have

$$\{\mathcal{V} | \text{valuation rings of } K(x) \text{ trivial on } K \text{ with } K(P) = K\} \cong \mathbf{P}(K).$$

The set $\{\mathcal{V}|\text{valuation rings of } K(x) \text{ trivial on } K\}$ is called the set of closed points of $\mathbf{P}$. They are associated with a maximal ideal of $K[x]$ or $K[x^{-1}]$. Of course, if $\mathcal{V}$ is associated with $(t(x))$ and $t(x) \neq x$, then the same $\mathcal{V}$ is associated with $x^{-\deg(t)}t(x) \in K[x^{-1}]$ because $v_P(x) = 0$ if $t(x) \neq x$. Thus we find

$$\{\text{closed points of } \mathbf{P}_{/K}\}$$
$$= \{\text{maximal ideals of } K[x]\} \cup \{\text{maximal ideals of } K[x^{-1}]\}.$$

We now treat a general case of algebraic function fields not necessarily a rational function field $K(x)$. For a given valuation v, the *residue field* of v means the residue field of the valuation ring of v. Let $\mathfrak{K}$ be an algebraic function field defined over K, and choose elements x and y so that $\mathfrak{K} = K(x, y)$ and y is separably integral over $K[x]$. Let R be the integral closure of $K[x]$ in $\mathfrak{K}$. Since $K[x]$ is a principal ideal domain and the quotient field of R (i.e., $\mathfrak{K}$) is a finite extension of $K(x)$, we find that R is free of rank $[\mathfrak{K} : K(x)]$ over $K[x]$. Since R is a finite integrally closed extension of the principal ideal domain $K[x]$, R is a Dedekind domain ([CRT] Section 11). So each nonzero fractional R-ideal $\mathfrak{a}$ has a prime decomposition $\mathfrak{a} = \prod_P P^{e(P)}$ into a product of prime ideals P ($e(P) \in \mathbb{Z}$). With each prime ideal of R, we can associate a valuation $v_P : \mathfrak{K} \to \mathbb{Z} \cup \{\infty\}$ by $v_P(f) = e(P)$ if $(f) = fR = \prod_P P^{e(P)}$. For any given valuation ring $\mathcal{V}$ associated with a valuation $v : \mathfrak{K} \to \mathbb{Z} \cup \{\infty\}$ with $v(K^\times) = 0$ but $v(\mathfrak{K}^\times) = \mathbb{Z}$, we find $x \in \mathfrak{K}$ transcendental over K such that $x \in \mathcal{V}$. Thus all valuations v as above are of the form v_P for a suitable choice of x and a prime ideal P of an integral closure R of $K[x]$. Thus we may identify

$$V(K) = \left\{ v\text{:discrete valuation} \middle| \begin{matrix} v(K^\times) = 0 \text{ and } v(\mathfrak{K}^\times) = \mathbb{Z} \\ \text{with residue field } K \end{matrix} \right\}.$$

Any prime ideal P of $K[x]$ is decomposed into a product of primes $P_1, \ldots, P_g$ in R. If K is algebraically closed, then there are g distinct points over a given point $P = (x - \alpha)$ of $\mathbf{P}$ for $\alpha \in K$. If K is not algebraically closed, $K(P_i)$ for some i may be a nontrivial finite extension of K. In other words, in the algebraic closure $\overline{K}$ of K, we have at most $[K(P_i) : K]$ (geometric) points coming out of P_i corresponding to embeddings of $K(P_i)$ into $\overline{K}$.

If furthermore, $R = K[x, y]$ for another element $y \in \mathfrak{K}$, we have an equation $f(X, Y) = Y^g + a_1(X)Y^{g-1} + \cdots + a_g(X)$ satisfied by (x, y). So for a given point $P = (x - \alpha) \in \mathbf{P}(K)$, the point $(\alpha, \beta) \in \overline{K}^2$ satisfying $f(\alpha, \beta) = 0$ gives rise to a point P_i over P in $\mathbf{P}^2$. In this way, $P_i \mapsto (\alpha, \beta)$ supplies us with a projective embedding of $V(\overline{K})$ into $\mathbf{P}^2(\overline{K})$. The image is the projective algebraic curve defined by the equation $f(X, Y) = 0$ in $\mathbf{P}^2$ (or more precisely, using the homogeneous coordinate (X, Y, Z), it is defined by $Z^g f(\frac{X}{Z}, \frac{Y}{Z}) = 0$). Thus except for finitely many points, there are really g distinct points in $V(\overline{K})$ over a given point $P \in \mathbf{P}(\overline{K})$. More generally, if $R = K[x, y_1, \ldots, y_m]$, then $P \mapsto (x, y_1, \ldots, y_m)$ gives an embedding of V into the $(m + 1)$-dimensional projective space $\mathbf{P}^{m+1}$.

If we take a proper subring $K[x,y] \subsetneq R$ such that $\mathfrak{K} = K(x,y)$, the ring $K[x,y]$ is not integrally closed. In other words, for $P \in V(K)$, x and y generate a subring $\mathcal{V}'_P$ of the valuation ring. Since $\mathfrak{m}_P \cap K[x] = (t(x))$ for an irreducible polynomial $t(x)$, $\mathfrak{m}_P^e \subset \mathcal{V}'_P$ for some $e \leq g$ (since $P^g \subset (t(x))\mathcal{V}_P$).

Let $f(X,Y)$ be the equation of (x,y). We can think of the projective algebraic curve $C \subset \mathbf{P}^2$ defined by this equation. Then if K is algebraically closed, $\alpha = (x \mod P) \in K(P) = K$ and $\beta = (y \mod P) \in K$ gives a point (α,β) of C in $\mathbf{P}^2$. Even if V does not ramify, in $C(K)$, $P' = (\alpha,\beta)$ has extra ramification, and $\mathcal{V}'_P$ is not a valuation ring.

We have a natural morphism $V \to C \to \mathbf{P}$ taking $P \mapsto (\alpha,\beta) \mapsto \alpha$. In this sense, V is the largest projective algebraic curve (over $\mathbf{P}$ with coordinate x) giving rise to the algebraic function field $\mathfrak{K}$.

For a projective algebraic curve C, we consider its function field $K(C)$ and $\mathcal{V}'_P$ made up of functions in $K(C)$ finite at $P \in C$. If $\mathcal{V}'_P$ is a valuation ring, we call P a smooth point. The projective algebraic curve C giving rise to a given algebraic function field $\mathfrak{K}$ is called a model of $\mathfrak{K}$. Among models of $\mathfrak{K}$, there is a unique model smooth everywhere, which is called the *smooth* or *non-singular* model of $\mathfrak{K}$. If $\mathfrak{K}'/\mathfrak{K}$ is a finite extension, by the above construction, we have a covering map $V' \to V$ as defined in the theorem. □

For each point $P \in V(K)$ for V as in the theorem, $f \in \mathfrak{K}$ has an expansion $f(t) = \sum_{n \gg -\infty} a_n t^n$ with $a_n \in K(P)$ for a generator t of $P\mathcal{V}_P$. An element $t \in \mathfrak{K}$ giving rise to the generator t of P is called a *uniformizer* at P.

Corollary 2.25 *Let the notation and assumption be as in the theorem. Then the morphism $\pi : V' \to V$ is a polynomial map of the projective coordinates.*

Proof. Let R (resp. R') be the integral closure of $K[x]$ in $\mathfrak{K}$ (resp. $\mathfrak{K}'$). Choose generators so that $R = K[x, y_1, \ldots, y_m]$ and $R' = K[x, y'_1, \ldots, y'_n]$. Since $R \subset R'$, we find $x = f_0(x, y'_1, \ldots, y'_n)$ and $y_i = f_i(x, y'_1, \ldots, y'_n)$. Then $\pi((x, y'_1, \ldots, y'_n)) = (x, y_1, \ldots, y_m)$, as desired. □

Since any field embedding $\sigma : \mathfrak{K} \to \mathfrak{K}'$ brings $\mathfrak{K}$ into a subfield $\sigma(\mathfrak{K})$, we have a morphism $\pi : V' \to V_\sigma$ of projective algebraic curves V' and V_σ corresponding to $\mathfrak{K}'$ and $\sigma(\mathfrak{K})$ as long as σ leaves the field of definition stable (even if σ is nontrivial on K). If $\mathfrak{K}$ is defined by $f(X, Y_1, \ldots, Y_m)$, then $\sigma(\mathfrak{K})$ is obviously defined by the polynomial $\sigma(f)(X, Y_1, \ldots, Y_m)$ obtained from f by applying σ to the coefficients in K of $f(X, Y_1, \ldots, Y_m)$. Thus V_σ as above is actually given by the conjugate $\sigma(V)$ (i.e., $\sigma(V)(\overline{K}) = \sigma(V(\overline{K}))$ for any choice of extension of σ to $\overline{K}$). We have proven:

Corollary 2.26 *Let $V_{/K}$ and $V'_{/K}$ be smooth projective curves. Then we have a canonical isomorphism $\mathrm{Hom}_K(V,V') \cong \mathrm{Hom}_K(K(V'),K(V))$ given by $K(V') \ni \phi \mapsto \phi \circ f$ for a morphism $f : V \to V'$ of projective curves, where $\mathrm{Hom}_K(V,V')$ is the collection of morphisms of projective curves defined over K and $\mathrm{Hom}_K(K(V'),K(V))$ is the set of all field homomorphisms over K.*

For each $f \in \mathrm{Hom}_K(V, V')$, taking the corresponding field homomorphism $\sigma : K(V') \hookrightarrow K(V)$, we define the degree of the morphism f by the field extension degree $[K(V) : \sigma(K(V'))]$. By definition, $\deg(f)$ also can be given by $\mathrm{rank}_{\mathcal{V}_{P'}} \left(\oplus_{P \in f^{-1}(P')} \mathcal{V}_P\right)$ for any choice of a closed point $P' \in V'$.

Exercises

1. Prove the surjectivity of v_P so that $\mathcal{V}_P$ in (2.6) is a DVR.
2. For a valuation ring $\mathcal{V}$ of $K(x)$ of a valuation trivial on K but nontrivial on $K(x)$, if $x \notin \mathcal{V}$, show that $\mathcal{V} = K[z]_{(z)}$ for $z = x^{-1}$.
3. Describe the intersection in $\mathbf{P}(K)$:

$$\{\text{maximal ideals of } K[x]\} \cap \{\text{maximal ideals of } K[x^{-1}]\}.$$

4. Give a more detailed proof of Corollary 2.25.

2.4.2 Zariski Topology

Let $\mathfrak{K}$ be an algebraic function field over an algebraically closed field K. We write V for its Zariski–Riemann space or equivalently the set of K-rational points of the associated smooth projective algebraic curve. We equip $V(K)$ with a topology in the following way. We cover V by "open" affine subvarieties. In the language of schemes, an open affine subvariety $U \subset V$ is given by the spectrum of an integral domain R; that is, $U = Spec(R)$ and $V - U$ is a proper closed subset of V (i.e., heuristically, a subset defined by the zero set of polynomial equations whose meaning becomes clear later).

Recall that v_P is the valuation $v_P : \mathfrak{K} \twoheadrightarrow \mathbb{Z} \cup \{\infty\}$ associated with each point $P \in V(K)$. We put

$$\begin{aligned} &\mathcal{V}_P = \left\{ f \in \mathfrak{K} \middle| v_P(f) \geq 0 \right\} \text{ (the discrete valuation ring of } v_P), \\ &\mathfrak{m}_P = \left\{ f \in \mathcal{V}_P \middle| v_P(f) > 0 \right\} \text{ (the maximal ideal of } \mathcal{V}_P), \text{ and} \\ &K(P) = \mathcal{V}_P/\mathfrak{m}_P \text{ (the residue field of } \mathcal{V}_P). \end{aligned} \tag{2.7}$$

For each nonconstant $x \in \mathfrak{K}$, we can think of

$$R_x = \bigcap_{\mathcal{V}_P \ni x} \mathcal{V}_P \subset \mathfrak{K}. \tag{2.8}$$

Then R_x is the integral closure of the polynomial ring $K[x]$ in $\mathfrak{K}$ (Exercise 1), and $\mathcal{V}_P$ is the localization of R_x at $P_x = P \cap R_x$ (Exercise 2). Then the set $Spec(R_x)(K)$ of all maximal ideals of R_x is in bijection with $V_x = \left\{ P \in V \middle| \mathcal{V}_P \ni x \right\}$ (see [GME] 1.2 and 1.4.3 for $Spec(R_x)$). We also have $V = \bigcup_{y \in \mathfrak{K} - K} V_y = V_x \cup V_{1/x}$. For each $P \in V$ and for $x \in \mathcal{V}_P$ (so, $P \in V_x$), V_x is called an open affine neighborhood of P. Any finite intersection $V_{x_1} \cap \cdots \cap V_{x_m}$ of $V_{x_j} \ni P$ $(j = 1, 2, \ldots, m)$ is also called an open affine neighborhood of P. Each open affine neighborhood U of x is in bijection with the set of maximal

ideals $Spec(R)(K)$ of a subring $R \subset \mathfrak{K}$ finitely generated over K. For example, $V_x \cong Spec(R_x)(K)$ and $V_x \cap V_y \cong Spec(R_x \cdot R_y)(K)$ ($R_x \cdot R_y$ is the composite ring of R_x and R_y in $\mathfrak{K}$), because $\mathcal{V}_P \supset R_x \cdot R_y \Leftrightarrow \mathcal{V}_P \supset R_x$ and $\mathcal{V}_P \supset R_y$. This topology we defined on V is called the *Zariski topology* on V.

2.4.3 Divisors

We now introduce divisors on an algebraic curve V. We start with interpreting elements in $\mathfrak{K}$ as a function defined on V with values in the projective space $\mathbf{P}^1$. Each $f \in \mathfrak{K} - K$ gives rise to an inclusion $K(f) \hookrightarrow \mathfrak{K}$ and hence gives rise to a (rational or meromorphic) function $f : V \to \mathbf{P}^1$ with

$$f(P) = \begin{cases} P \cap K[f] \in \mathbf{P}^1_f & \text{if } P \in V_f, \\ P \cap K[\frac{1}{f}] \in \mathbf{P}^1_{1/f} & \text{if } P \in V_{1/f}, \end{cases}$$

where $\mathbf{P}^1_g = \{P \in \mathbf{P}^1 | P \ni g\}$. This is well-defined by the following. If $P \in V_f$, then $P \cap K[f] = (f - \alpha)$ for $\alpha \in K = \mathbf{P}^1_f$, and $\alpha = (f \mod P) = f(P)$. If $P \in V_f \cap V_{1/f}$, then $P \cap K[\frac{1}{f}] = (\frac{1}{f} - \frac{1}{\alpha})$ and still $f(P) = \alpha$. If $P \notin V_f$, then $f(P) = \infty$. When $f(P) = 0$, we call P a zero of f and when $f(P) = \infty$, we call P a pole of f. For $f \in K$, we associate a constant function $f : V \to \mathbf{P}^1$ with value f everywhere. In this way, we can identify $\mathfrak{K}$ with the set of all meromorphic functions on V.

For each $P \in V$, take $t = t_P$ with $v_P(t_P) = 1$ (such a t_P always exists because $v_P : \mathfrak{K} \to \mathbb{Z} \cup \{\infty\}$ is surjective). By definition, $\mathfrak{m}_P = (t_P)$ in $\mathcal{V}_P$. Then we can expand $f \in \mathfrak{K}$ into a Laurent series of t_P as follows. First suppose that P is not a pole of f. Then $f - f(P)$ has zero at P; so, $f_1 = (f - f(P))/t_P$ does not have a pole at P; so, we have $f - (f(P) + f_1(P)t_P)$ is divisible by t_P^2 and so on. We thus have $f = \sum_{n \geq 0} c_n t_P^n \in K[[t_P]]$. If f has a pole at P, killing the pole by multiplying f by $t_P^{|v_P(f)|}$, $g = t_P^{|v_P(f)|} f$ does not have a pole at P; so, we apply the above argument, and after expanding g into a power series of t_P, we divide the power series by $t_P^{|v_P(f)|}$ to get the Laurent expansion of f. We have $f = \sum_{n \geq v_P(f)} c_n t_P^n \in K((t_P)) = K[[t_P]][\frac{1}{t_P}]$. If we choose another t'_P with $v_P(t'_P) = 1$, we find that $t_P = t'_P(c_0 + c_1 t'_P + c_2 {t'_P}^2 + \cdots)$ with $c_0 \neq 0$. Thus $t'_P/t_P \in \mathcal{V}_P^\times$ and if we write the power series expansion of f with respect to t_P as $f(t_P)$, then we get $f(t'_P)$ by substituting $t_P(t'_P)$ for t_P (check that this substitution gives a well-defined power series in t'_P because $t_P(P) = 0$).

The divisor $\mathrm{div}(f)$ of $f \neq 0$ is defined by a formal sum $\sum_P v_P(f)P$ (in some classical books in number theory, they use multiplicative notation, such as $\mathrm{div}(f) = \prod_P P^{v_P(f)}$, because it corresponds to the prime decomposition of the ideal (f), but we use additive symbols following the geometric tradition). Since f satisfies a polynomial equation of finite degree over $K[x]$ for a suitably chosen $x \in \mathfrak{K}$, f has only finitely many zeros and poles; thus the sum defining $\mathrm{div}(f)$ is actually a finite sum (because $v_P(f) = 0$ except for finitely many points $P \in V$).

We define the *divisor* group $\mathrm{Div}(V) = \mathrm{Div}(\mathfrak{K})$ by the free abelian group of all formal finite sums $D = \sum_{P \in V} e_P P$, where $e_P = 0$ except for finitely many $P \in V$. We define the *degree* of a divisor $D = \sum_{P \in V} e_P P$ by $\deg(D) = \sum_P e_P$. We have a homomorphism $\deg : \mathrm{Div}(V) \to \mathbb{Z}$, and we define $\mathrm{Div}^0(V) = \mathrm{Ker}(\deg)$.

Proposition 2.27 *If $f \in \mathfrak{K}$, then $\deg(\mathrm{div}(f)) = 0$; so, $\mathrm{div}(f) \in \mathrm{Div}^0(V)$.*

Proof. We may assume that $f \in \mathfrak{K} - K$, because the assertion is clear for $f \in K^\times$. The inclusion $K(f) \hookrightarrow \mathfrak{K}$ induces a projection $f : V \to \mathbf{P}^1$. Then for each $p \in \mathbf{P}^1$, regarding it as a prime ideal of $K[f]$, we find $pR = \prod_P P^{e_P}$ with prime ideals P in the integral closure R of $K[f]$. Then by definition, we have $f(P) = 0 \Leftrightarrow e_P > 0$ and $\sum_P e_P = [R : K[f]] = [\mathfrak{K} : K(f)]$. In particular, writing $\infty = \prod_Q Q^{e'_Q}$ for the points Q with $f(Q) = \infty$, we find that

$$\deg(\mathrm{div}(f)) = \sum_P e_P - \sum_Q e'_Q = [\mathfrak{K} : K(f)] - [\mathfrak{K} : K(f)] = 0$$

as desired. □

Write $\mathcal{P}(V) = \{\mathrm{div}(f) | f \in \mathfrak{K}^\times\}$. Then $\mathcal{P}(V)$ is a subgroup of $\mathrm{Div}^0(V)$. The quotient group $\mathrm{Pic}(V) = \mathrm{Div}(V)/\mathcal{P}(V)$ (resp. $\mathrm{Jac}(V) = \mathrm{Pic}^0(V) = \mathrm{Div}^0(V)/\mathcal{P}(V)$) is called the Picard group (resp. the Jacobian) of V. When $D - D' \in \mathcal{P}(V)$, we say that D is *linearly* equivalent to D', and if $\deg(D) = \deg(D')$, we call D is *algebraically* equivalent to D'.

2.4.4 Differentials

We introduce a notion of differential forms on an algebraic curve V in a purely algebraic way. For the moment, until the end of the proof of Theorem 2.29, we suppose that K is algebraically closed.

For a given $f \in \mathfrak{K}$, we formally define df as a collection of power series $df(t_P) = \frac{df(t_P)}{dt_P} dt_P$. Here P runs through all points $P \in V$, and

$$\frac{d(\sum_n c_n t_P^n)}{dt_P} dt_P = \left(\sum_n c_n n t_P^{n-1} \right) dt_P.$$

In particular, df does not have the term $\frac{1}{t_P}$. If we change coordinate t_P into t'_P, the expression of df will change according to the chain rule:

$$df(t'_P) = \frac{df(t'_P)}{dt'_P} dt'_P = \frac{df(t'_P(t_P))}{dt'_P} \frac{dt'_P}{dt_P} dt_P.$$

We also think of $\omega = gdf$ as a collection of power series $\{g(t_P) \frac{df(t_P)}{dt_P} dt_P\}_{P \in V}$. In other words, if gdf and $g'df'$ give rise to the same power series at every $P \in V$, we identify $gdf = g'df'$. The totality of all differentials on V gives

rise to a vector space over $\mathfrak{K}$. We see $gdf/g'df' = \frac{g}{g'}\frac{df}{df'} \in \mathfrak{K}$, because f and f' satisfy the polynomial relation $F(f,f') = 0$ over K. In other words, if $F(X,Y) = \sum_{i,j} a_{i,j}X^iY^j$, then taking a formal derivative of $F(f,f') = 0$, we have

$$0 = dF(f,f') = \sum_{i,j}\left(ia_{i,j}f^{i-1}f'^{j}df + ja_{i,j}f^{i}f'^{j-1}df'\right),$$

and df and df' are linearly dependent over $\mathfrak{K}$. Thus this vector space over $\mathfrak{K}$ is one-dimensional.

We define $v_P(gdf) = v_P(g\frac{df(t_P)}{dt_P})$. Note here that $v_P\left(\frac{dt'_P}{dt_P}\right) = 0$ for two parameters t_P and t'_P at $P \in V$. By the chain rule, $\frac{df}{dt_P} = \frac{df}{dt'_P}\frac{dt'_P}{dt_P}$, we have

$$v_P\left(g\frac{df}{dt_P}\right) = v_P\left(g\frac{df}{dt'_P}\frac{dt'_P}{dt_P}\right) = v_P\left(g\frac{df}{dt'_P}\right) + v_P\left(\frac{dt'_P}{dt_P}\right) = v_P\left(g\frac{df}{dt'_P}\right).$$

Since the last expression of the above formula is the value $v_P(gdf)$ computed with respect to the coordinate t'_P, the value $v_P(gdf)$ is well-defined independently of the choice of the parameter t_P around P. In particular, we have

$$gdf = \left(g\frac{df(t_P)}{dt_P}\right)dt_P = \left(\sum_{n \geq v_P(gdf)} c_n t_P^n\right)dt_P.$$

We define $\operatorname{div}(\omega) = \sum_P v_P(\omega)P$.

Since differentials form a one-dimensional vector space over $\mathfrak{K}$, for any two differentials ω and ω', $\operatorname{div}(\omega)$ and $\operatorname{div}(\omega')$ are linearly equivalent. Thus $\deg(\operatorname{div}(\omega))$ is independent of the choice of ω. Writing $\deg(\omega) = 2g-2$, we define the genus $g = g(V)$ of V.

A differential ω is called *holomorphic* or *of the first kind* if $v_P(\omega) \geq 0$ for all $P \in V$. If ω' is holomorphic, writing $\omega' = f\omega$, we find $\operatorname{div}(f) + \operatorname{div}(\omega) = \operatorname{div}(\omega') \geq 0$ (we write $D \geq 0$ for $D = \sum_P e_P P$ if $e_P \geq 0$ for all P). We have

$$f \in L(\operatorname{div}(\omega)) = \left\{g \in \mathfrak{K} \middle| \operatorname{div}(g) \geq -\operatorname{div}(\omega)\right\}.$$

The space $\Omega_{V/K}$ of all holomorphic differentials is isomorphic to $L(\operatorname{div}(\omega))$ by $f\omega \leftrightarrow f$ for any choice ω of nonzero differentials, and $\dim_K L(\operatorname{div}(\omega)) = \dim_K \Omega_{V/K}$ is finite and independent of the choice of ω (see Lemma 2.30). We later show (see Theorem 2.31) that this dimension is given by the genus g; so, $g \geq 0$.

For a differential $\omega = gdf$, we expand it into a power series in t_P at $P \in V$, and write

$$\omega = \left(\sum_{n \geq v_P(\omega)} c_n t_P^n\right)dt_P.$$

Then we define $\operatorname{Res}_P(\omega) = c_{-1}$.

Proposition 2.28 *The residue* $\mathrm{Res}_P(\omega)$ *is well-defined independently of the choice of the parameter* t_P.

Since ω has only finitely many poles, $\mathrm{Res}(\omega) = \sum_P \mathrm{Res}_P(\omega) \in K$ is well-defined.

Proof. We give a proof valid only when K is of characteristic 0. See [ALF] Appendix and [ALG] Theorems III.7.14.1–2 for the proof valid for arbitrary characteristics. Choose two parameters $t = t_P$ and $t' = t'_P$, and expand $t' = a_1 t + a_2 t^2 + \cdots$ with $a_1 \neq 0$. If the expansion of $\omega = g df$ with respect to t is given by $g\frac{df}{dt'}dt' = \sum_n c_n t'^n dt'$, the one with respect to t is given by

$$g\frac{df}{dt}dt = g\frac{df}{dt'}\frac{dt'}{dt}dt = \sum_n c_n(a_1 t + a_2 t^2 + \cdots)^n \left(\sum_j j a_j t^{j-1}\right) dt.$$

Let us look into the term involving t^{-1}:

$$c_{-1}(a_1 t + a_2 t^2 + \cdots)^{-1}\left(\sum_{j\geq 1} j a_j t^{j-1}\right) dt = \frac{c_{-1}a_1}{a_1} + \text{higher terms}.$$

We need to show that $c_n(a_1 t + a_2 t^2 + \cdots)^n \left(\sum_j j a_j t^{j-1}\right) = c_n t'^n \frac{dt'}{dt}$ for $n \neq -1$ does not involve the term t^{-1}. This follows from $t'^n \frac{dt'}{dt} = \frac{1}{n+1}\frac{dt'^{n+1}}{dt}$, since the Laurent series expansion of $\frac{d\phi}{dt}$ does not involve the term t^{-1} which is not a derivative of a power of t. This finishes the proof.

Since $\mathfrak{K}/K(x)$ is a separable finite extension for a nonconstant $x \in \mathfrak{K}$, we have the trace map $\mathrm{Tr}_{\mathfrak{K}/K(x)} : \mathfrak{K} \to K(x)$ for any nonconstant function x. We study how the residue map behaves under field extensions using the trace map. The inclusion $K(x) \hookrightarrow \mathfrak{K}$ is induced by the projection $x : V \to \mathbf{P}^1$. For a prime ideal p of $K[x]$, we consider $x^{-1}(p) = \{P\}$; thus, $pR = \prod_P P^{e_P}$ for the integral closure R of $K[x]$ in $\mathfrak{K}$.

By the Chinese remainder theorem, $R/pR \cong \prod_P R/P^{e_P}$ and $R/p^n R \cong \prod_P R/P^{ne_P}$. Thus the completion of $\mathfrak{K}$ with respect to p is given by $\mathfrak{K}_p = \prod_{P:x(P)=p} \mathfrak{K}_P$. In particular, we have $\mathrm{Tr}_{\mathfrak{K}/K(x)}(f) = \sum_{P:x(P)=p} \mathrm{Tr}_{\mathfrak{K}_P/K(x)_p}(f)$ for $f \in \mathfrak{K}$, because $\mathrm{Tr}_{\mathfrak{K}/K(x)}(f)$ is $\mathrm{Tr}(\rho(f))$ for the matrix $\rho(f) \in M_n(K(x))$ given by $(fg_1, fg_2, \ldots, fg_d) = (g_1, \ldots, g_d)\rho(f)$ for a base g_j of $\mathfrak{K}/K(x)$.

In $\mathfrak{K}_P = K[[t]]$ $(t = t_P)$, we find $t^e = u t_p$ $(e = e_P)$ for a unit power series $u = c_0 + c_1 t + c_2 t^2 + \cdots$ with $c_0 \neq 0$. We can take an eth root v of u in $K[[t]]$ because K is algebraically closed. Changing t by vt, we may assume that $K(x)_p = K((t^{e_P}))$ and $t_p = t^e$. Then we realize that $K((t))$ is a Galois extension of $K(x)_p = K((t^e))$ with the Galois group isomorphic to the group of eth roots of unity μ_e. Each $\zeta \in \mu_e$ acts on t by $t \mapsto \zeta t$. Then

$$\mathrm{Tr}_{\mathfrak{K}_P/K(x)_p} t^n = \left(\sum_{\zeta\in\mu_e} \zeta^n\right) t^n = \begin{cases} 0 & e \nmid n, \\ e t^n & \text{if } e|n. \end{cases}$$

We see that for $t^n \in \mathfrak{K}_P$, $\mathrm{Res}_p(\mathrm{Tr}_{\mathfrak{K}_P/K(x)_p}(t^n)d(t^e)) = e$ or 0 accordingly as $n = -e$ or not. On the other hand, we see $\mathrm{Res}_P(t^n d(t^e)) = \mathrm{Res}_P(et^{n+e-1}dt) = e$ or 0 accordingly as $n = -e$ or not, because $t^n \frac{dt^e}{dt} = et^{n+e-1}$. Thus $\mathrm{Res}_p(\mathrm{Tr}_{\mathfrak{K}/K(x)}(f)dt_p) = \sum_{P:x(P)=p} \mathrm{Res}_P(f dt_P)$, and hence we get, for all $f \in \mathfrak{K}$,

$$\sum_{p \in \mathbf{P}^1} \mathrm{Res}_p(\mathrm{Tr}_{\mathfrak{K}/K(x)}(f)dx) = \sum_{P \in V} \mathrm{Res}_P(f dx). \tag{2.9}$$

Theorem 2.29 (Residue Theorem) *For every differential ω on V, we have $\sum_{P \in V} \mathrm{Res}_P(\omega) = 0$.*

Proof. Pick $x \in \mathfrak{K} - K$. Since the space of differentials is one-dimensional over $\mathfrak{K}$, we can write $\omega = f dx$ for $f \in \mathfrak{K}$. Then by (2.9), we may assume that $\mathfrak{K} = K(x)$. Since $\mathbf{P}^1 = K \cup \{\infty\}$, for each $\alpha \in K$, we take $t_\alpha = x - \alpha$ to be our parameter. At ∞, we choose $t_\infty = \frac{1}{x}$ to be the parameter there. For each $f \in K(x)$, expand $f = \sum_n c_n t_\alpha^n$ and define $[f]_\alpha = \sum_{n<0} c_n t_\alpha^n$. Then $[f]_\alpha = 0$ except for finitely many α and $\phi = f - \sum_{\alpha \in \mathbf{P}^1} [f]_\alpha \in K$, because $f - [f]_\alpha \in K[x]_{(x-\alpha)} = \mathcal{V}_\alpha$ and $\sum_{\beta \neq \alpha} [f]_\beta \in \mathcal{V}_\alpha$ (which implies the difference ϕ is a rational function without a pole; so, constant by $\deg(\phi) = 0$). We expand $f = \sum_{\alpha \in K, j} \frac{c_{\alpha,j}}{(x-\alpha)^j} + \sum_{j \geq 0} a_j x^j$. We only need to check the vanishing for $t_\alpha^{-j} dx = t_\alpha^{-j} dt_\alpha$ and $x^j dx$ for $j \geq 1$. By computation, we have $\mathrm{Res}_\alpha(x^j dx) = 0$ for all $\alpha \in K \cup \{\infty\}$ because $x^j dx = -t_\infty^{-j-2} dt_\infty$. Similarly (Exercise 5),

$$\mathrm{Res}_\beta(t_\alpha^{-j} dt_\alpha) = \begin{cases} 1 & \text{if } \beta = \alpha \text{ and } j = 1, \\ -1 & \text{if } \beta = \infty \text{ and } j = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{2.10}$$

This finishes the proof.

We now treat the general case; so, K is an arbitrary field inside its algebraic closure $\overline{K}$. We suppose that K is algebraically closed in $\mathfrak{K}$. The composite $\overline{K}\mathfrak{K} \cong \mathfrak{K} \otimes_K \overline{K}$ is an algebraic function field over $\overline{K}$. We write V (resp. $\overline{V}$) for the Zariski–Riemann space of $\mathfrak{K} = K(V)$ (resp. $\overline{K}\mathfrak{K} = \overline{K}(\overline{V})$). Then we have a natural inclusion $V(K) \hookrightarrow \overline{V}(\overline{K})$. Note here that $V(K) = \{P \in V | K(P) = K\}$. If $K(P) \supsetneq K$, then there are $[K(P) : K]$ points corresponding to P. In other words, for points $P_1, \ldots, P_d$ with $d = [K(P) : K]$ in $\overline{V}(\overline{K})$, $P_j \cap \mathfrak{K} = P \Leftrightarrow v_{P_j}|_{\mathfrak{K}} = v_P$ or, equivalently, P_j corresponds to each field embedding $K(P) \hookrightarrow \overline{K}$. We define $\mathrm{Res}_P(\omega) = \sum_j \mathrm{Res}_{P_j}(\overline{\omega})$, where $\overline{\omega}$ is the differential of $\overline{K}\mathfrak{K}$ corresponding to $\omega = g df$. Since $\mathrm{Res}_{P_j}(\overline{\omega})$ for different j are conjugates of each other, we find $\mathrm{Res}_P(\omega) \in K$. Similarly, we send P to the sum $\sum_j P_j$ (trace of P_j) in $\mathrm{Div}(\overline{V})$ and in this way, we embed $\mathrm{Div}(V)$ into $\mathrm{Div}(\overline{V})$. This embedding sends $\mathcal{P}(V)$ into $\mathcal{P}(\overline{V})$, and we have deg on $\mathrm{Div}(V)$ by pulling back deg defined on $\mathrm{Div}(\overline{V})$. In this way, everything we stated (in

particular, the residue theorem, Theorem 2.29) is valid over the general field K (not necessarily algebraically closed).

Similarly to $L(\mathrm{div}(\omega))$, we define for a general divisor D,

$$L(D) = \{f \in \mathfrak{K} | \, \mathrm{div}(f) \geq -D\},$$

where for two divisors $D = \sum_P e_P P$ and $D' = \sum_P e'_P P$, we write $D \geq D'$ if $e_P \geq e'_P$ for all P. Plainly $L(D)$ is a K-vector space.

Lemma 2.30 *The dimension* $\dim_K L(D)$ *is finite. In particular, we have* $L(0) = K$ *and* $\dim_K L(0) = 1$.

Proof. If f is nonconstant, $\mathrm{div}(f)$ is nontrivial (in $\overline{V}$) and $\deg(\mathrm{div}(f)) = 0$. Thus $L(0) \subset \overline{K} \cap \mathfrak{K} = K$ (by the definition of algebraic function field, the integral closure in $\mathfrak{K}$ of the base field is itself). For sufficiently positive D', we see $L(D) \subset L(D')$. Writing $D' = \sum_P e'_P P$ and expanding $f \in L(D')$ into a Laurent series of t_P at $P \in D'$, we have e'_P linear forms taking $f \in L(D')$ to the coefficient of t_P^{-j} for $0 < j \leq e'_P$. If all such linear forms vanish at f, we see $f \in L(0)$ and hence f is a constant. Thus $\dim_K L(D') < \infty$, which proves the desired assertion. □

Since $f \cdot L(D) = L(D - \mathrm{div}(f))$, the dimension $\dim_K L(D)$ depends only on the linear equivalence class of D. For the linear equivalence class Ω of differential divisors, we have a well-defined number $\dim_K L(\Omega - D)$.

We introduce the following Riemann–Roch theorem (see [BNT] Chapter VI, [FAN] 7.2 and [GME] 2.1.3 for different proofs).

Theorem 2.31 *We have for all* $D \in \mathrm{Div}(V)$,

$$\dim_K L(D) = \deg(D) - g + 1 + \dim_K L(\Omega - D),$$

where g *is the genus of the curve* V.

Applying this theorem to the trivial divisor 0, we find

$$1 = \dim_K L(0) = \deg(0) - g + 1 + \dim_K L(\Omega),$$

which shows that $g = \dim_K L(\Omega) = \dim_K \Omega_{V/K}$ as we claimed before.

Exercises

1. Prove that R_x in (2.8) is the integral closure of $K[x]$ in $\mathfrak{K}$.
2. Prove that $\mathcal{V}_P$ is the localization of R_x at $P_x = P \cap R_x$.
3. Show that V is not a Hausdorff space under its Zariski topology.
4. Prove Theorem 2.28 for an algebraic function field of characteristic $p > 0$.
5. Give a detailed proof of (2.10).

2.4.5 Adele Rings of Algebraic Function Fields

We insert a brief interpretation of differentials by the language of adeles (due to A. Weil; see [ALF] Chapter 2). We write $\mathfrak{K}_P$ for the Laurent series ring: $K(P)((t_P))$ and consider the product ring $\prod_{P\in V}\mathfrak{K}_P$, which we consider as a $\mathfrak{K}$-algebra by embedding $\mathfrak{K}$ diagonally. We define the adele ring $\mathfrak{K}_{\mathbb{A}}$ of $\mathfrak{K}$ by

$$\mathfrak{K}_{\mathbb{A}} = \mathfrak{K} + \widehat{\mathcal{V}} \subset \prod_{P\in V}\mathfrak{K}_P,$$

where $\widehat{\mathcal{V}} = \prod_{P\in V}\mathcal{V}_P$ for $\mathcal{V}_P = K(P)[[t_P]]$. Here we do not assume that K is algebraically closed. Similarly to the case of $\mathbb{Q}$, we can identify

$$\mathfrak{K}_{\mathbb{A}} = \left\{(f_P)_{P\in V} \middle| f_P \in \mathcal{V}_P \ \text{ except for finitely many } P\right\}. \tag{2.11}$$

For a divisor $D = \sum_P e_P P$, we define $\widehat{\mathcal{V}}(D) = \{(f_P) \in \widehat{\mathcal{V}} | v_P(f_P) \geq -e_P\}$. We equip with $\mathfrak{K}_{\mathbb{A}}$ the topology whose system of neighborhoods of $f \in \mathfrak{K}_{\mathbb{A}}$ is given by $f+\widehat{\mathcal{V}}(D)$ for all negative divisors D. Here we call D negative if $e_P \leq 0$ for all P. By definition, $\mathfrak{K}_{\mathbb{A}}$ becomes a topological ring. For each $f = (f_P)_P \in \mathfrak{K}_{\mathbb{A}}^{\times}$, we define its divisor $\mathrm{div}(f) = \sum_P v_P(f_P)P$. This is well-defined since $v_P(f_P) = 0$ for almost all P if $f \in \mathfrak{K}_{\mathbb{A}}^{\times}$. In particular, $\widehat{\mathcal{V}}^{\times} = \{f \in \mathfrak{K}_{\mathbb{A}} | \mathrm{div}(f) = 0\}$. Thus we have $\mathrm{Div}(\mathfrak{K}_{\mathbb{A}}^{\times}) = \mathrm{Div}(V)$ and hence

$$\mathrm{Pic}(V) = \mathfrak{K}_{\mathbb{A}}^{\times}/\mathfrak{K}^{\times}\widehat{\mathcal{V}}^{\times}. \tag{2.12}$$

For each differential ω on V (supposing K to be algebraically closed), we define a K-linear form $\widetilde{\omega} : \mathfrak{K}_{\mathbb{A}} \to K$ by $f \mapsto \sum_P \mathrm{Res}_P(f\omega)$. Since $f_P \in \mathcal{V}_P$ for almost all P, $\mathrm{Res}_P(f\omega) = 0$ for almost all P; so, the linear form $\widetilde{\omega}$ is well-defined. By definition, if $\mathrm{div}(f\omega) = \mathrm{div}(f) + \mathrm{div}(\omega) \geq 0$, then $\widetilde{\omega}(f) = 0$. Thus $\widetilde{\omega}(\widehat{\mathcal{V}}(\mathrm{div}(\omega))) = 0$, and hence $\widetilde{\omega}$ is a continuous linear form on $\mathfrak{K}_{\mathbb{A}}$ with respect to the discrete topology on K and vanishes over (the diagonal image) $\mathfrak{K}$ by Theorem 2.29.

We consider the space $\mathcal{L}$ of continuous linear forms $f : \mathfrak{K}_{\mathbb{A}}/\mathfrak{K} \to K$. We know that $\mathcal{L}$ is a vector space over $\mathfrak{K}$. Actually $\mathcal{L}$ is isomorphic to the space of differentials by $\omega \mapsto \widetilde{\omega}$. We could have defined the differentials as K-linear forms on $\mathfrak{K}_{\mathbb{A}}/\mathfrak{K}$, and the proof of the Riemann–Roch theorem can be given from this point of view (see [BNT] Chapter VI and [ALF] Chapter 2).

Exercises

1. Give a detailed proof of (2.11).
2. Prove that $\mathfrak{K}_{\mathbb{A}}$ is locally compact if and only if K is a finite field.
3. Prove that $\widehat{\mathcal{V}}(D) + \widehat{\mathcal{V}}(D') = \widehat{\mathcal{V}}([D, D'])$ for the least common multiple $[D, D']$ of D and D'.
4. Prove that $\widehat{\mathcal{V}}(D) \cap \widehat{\mathcal{V}}(D') = \widehat{\mathcal{V}}((D, D'))$ for the greatest common divisor (D, D') of D and D'.

2.5 Elliptic Curves over a Field

An algebraic function field $\mathfrak{K}/K$ is called *elliptic* if its genus is equal to 1. The corresponding algebraic curve E/K with one designated point $\mathbf{0} \in E(K)$ is called an elliptic curve. The elliptic curve is a pair $(E, \mathbf{0})/K$. The point $\mathbf{0}$ is called the origin of E, and two elliptic curves E and E' are isomorphic if we have an isomorphism $\phi : E \cong E'$ of algebraic curves that sends the origin to the origin. We study elliptic curves over a field in detail here. When we regard the point $\mathbf{0}$ as a divisor, we write $[\mathbf{0}]$ instead of $\mathbf{0}$.

2.5.1 Dimension Formulas

For divisors $m[\mathbf{0}]$, we study $L(m[\mathbf{0}])$. Since $0[\mathbf{0}]$ is the trivial divisor 0, we find $L(0[\mathbf{0}]) = K$. By definition, $\Omega_{E/K}$ is one-dimensional over K. Since $\deg(\omega) = 2g - 2 = 0$ for a differential ω on E, if ω has a zero, then ω has to have a pole. If ω is holomorphic nonzero, it vanishes nowhere. In other words, nowhere-vanishing differentials are all nonzero multiples of ω. In particular, $\operatorname{div}(\omega) = 0$. Recall the linear equivalence class of $\operatorname{div}(\omega')$ for all meromorphic differentials ω' on E. Then Ω is the linear equivalence class of all meromorphic functions in $\mathfrak{K}$, and $L(\Omega - D) \cong L(-D)$. If $D > 0$, then $\dim_K L(-D) = 0$ (Exercise 1).

We fix a nowhere-vanishing differential ω. If $f \in L([\mathbf{0}])$, then f has possibly only one simple pole at $\mathbf{0}$. Thus $\operatorname{Res}_{\mathbf{0}}(f\omega) = \sum_{P \in E} \operatorname{Res}_P(f\omega) = 0$ by Theorem 2.29. The function f is in $L(0)$ and is constant, and $\dim_K L([\mathbf{0}]) = \dim_K L(0) = 1$. By the Riemann–Roch theorem, we find

$$\dim_K L(m[\mathbf{0}]) = \deg(m[\mathbf{0}]) - g + 1 + \dim_K L(-m[\mathbf{0}]) = m \quad \text{if } m > 0. \tag{2.13}$$

More generally, by the same proof, we get

Proposition 2.32 *Let E be an elliptic curve over a field K. If $D > 0$ is a positive divisor, we have* $\dim_K L(D) = \deg(D)$.

We may normalize the uniformizing parameter at the origin $\mathbf{0}$ in terms of a holomorphic differential ω. Choose a parameter t at the origin. Since ω does not vanish at $\mathbf{0}$, we have $\omega = (c_0 + c_1 t + c_2 t^2 + \cdots)dt$ with $c_0 \neq 0$. Making a variable change $t = c_0^{-1}T$, we get $\omega = (1 + a_1 T + a_2 T^2 + \cdots)dT$. The parameter T with the above property is unique modulo T^2 in $K[[T]]$. Conversely, once we have chosen a formal parameter T modulo T^2, there is a unique nowhere-vanishing differential ω of the above form because of $\dim_K \Omega_{E/K} = 1$ (Exercise 2). Such a T is called "*T adapted to ω.*"

Exercises

1. Show that $\dim_K L(-D) = 0$ for an elliptic curve E if $D > 0$.
2. Show that for a given parameter t at $\mathbf{0}$, there exists a unique nowhere-vanishing differential ω so that its expansion in t has constant term 1.

2.5.2 Weierstrass Equations of Elliptic Curves

We now embed $E_{/K}$ into the two-dimensional projective space $\mathbf{P}^2_{/K}$ using a base of $L(3[\mathbf{0}])$ and determine the equation of the image in $\mathbf{P}^2_{/K}$. We first consider $L(n[\mathbf{0}])$ which has dimension n if $n > 0$. We have $L([\mathbf{0}]) = K$ and $L(2[\mathbf{0}]) = K1 + Kx$. Since x has to have a pole of order 2 at $\mathbf{0}$, we may normalize x so that $x = T^{-2}(1 + \text{higher terms})$ in $K[[T]]$. Here x is unique up to translation: $x \mapsto x + a$ with $a \in K$. Then $L(3[\mathbf{0}]) = K1 + Kx + Ky$. We may then normalize y so that $y = -T^{-3}(1 + \text{higher terms})$ (following the tradition, we later rewrite y for $2y$; thus, the normalization will be $y = -2T^{-3}(1 + \text{higher terms})$ at the end). Then y is unique up to the affine transformation: $y \mapsto y + ax + b$ $(a, b \in K)$.

Proposition 2.33 *Suppose that the characteristic of the base field K is different from 2 and 3. Then for a given pair (E, ω) of an elliptic curve E and a nowhere-vanishing differential ω both defined over K, we can find a unique base $(1, x, y)$ of $L(3[\mathbf{0}])$ such that E is embedded into $\mathbf{P}^2_{/K}$ by $(1, x, y)$ whose image is defined by the affine equation*

$$y^2 = 4x^3 - g_2 x - g_3 \quad \textit{with} \quad g_2, g_3 \in K, \tag{2.14}$$

and ω on the image is given by $\frac{dx}{y}$. Conversely, a projective algebraic curve defined by the above equation is an elliptic curve with a specific nowhere-vanishing differential $\frac{dx}{y}$ if and only if the discriminant $\Delta(E, \omega) = g_2^3 - 27g_3^2$ of $4X^3 - g_2 X - g_3$ does not vanish.

An equation of an elliptic curve E as in (2.14) is called a *Weierstrass equation* of E, which is determined by the pair (E, ω).

Proof. By the dimension formulas, counting the order of poles at $\mathbf{0}$ of monomials of x and y, we have

$$\begin{aligned} L(4[\mathbf{0}]) &= K + Kx + Ky + Kx^2, \\ L(5[\mathbf{0}]) &= K + Kx + Ky + Kx^2 + Kxy \quad \text{and} \\ L(6[\mathbf{0}]) &= K + Kx + Ky + Kx^2 + Kxy + Kx^3 \\ &= K + Kx + Ky + Kx^2 + Kxy + Ky^2, \end{aligned}$$

from which the following relation results,

$$y^2 + a_1 xy + a_3 y = x^3 + a_2 x^2 + a_4 x + a_6 \quad \text{with } a_\jmath \in K, \tag{2.15}$$

because the poles of order 6 of y^2 and x^3 have to be canceled. Therefore $E_{/K}$ can be embedded into $\mathbf{P}^2_{/K}$ via $P \mapsto (x(P), y(P))$. The image is defined by the equation (2.15).

Now we make a variable change $y \mapsto y + ax + b$ in order to remove the terms of xy and y (i.e., we are going to make $a_1 = a_3 = 0$):

$$(y+ax+b)^2+a_1x(y+ax+b)+a_3(y+ax+b)$$
$$=y^2+(2a+a_1)xy+(2b+a_3)y+\text{polynomial in } x.$$

Assuming that 2 is invertible in K, we take $a=-\frac{a_1}{2}$ and $b=-\frac{a_3}{2}$. The resulting equation is of the form $y^2=x^3+b_2x^2+b_4x+b_6$. We now make the change of variable $x\mapsto x+a'$ to make $b_2=0$:

$$y^2=(x+a')^3+b_2(x+a')^2+b_4(x+a')+b_6=x^3+(3a'+b_2)x^2+\cdots.$$

Assuming that 3 is invertible in K, we take $a'=-\frac{b_2}{3}$. We can rewrite the equation as in (2.14) (making a variable change $2y\mapsto y$). By the variable change as above, we have $y=-2T^{-3}(1+\text{higher terms})$, and from this, we conclude $\omega=\frac{dx}{y}$. The numbers g_2 and g_3 are determined by T adapted to a given nowhere-vanishing differential form ω.

Conversely, we have seen that any curve defined by equation (2.14) is smooth in Example 2.23 if the cubic polynomial $F(X)=4X^3-g_2X-g_3$ has three distinct roots in K. In other words, if the discriminant $\Delta(E,\omega)$ of $F(X)$ does not vanish, E is smooth.

For a given equation, $Y^2=F(X)$, the algebraic curve E defined by the homogeneous equation $Y^2Z=4X^3-g_2XZ^2-g_3Z^3$ in $\mathbf{P}^2_{/K}$ has a rational point $\mathbf{0}=(0,1,0)\in E(K)$, which is ∞ in $\mathbf{P}^2$. Thus E is smooth over K if and only if $\Delta(E,\omega)\neq 0$ (Exercise 2).

We show that there is a canonical nowhere-vanishing differential $\omega\in\Omega_{E/K}$ if E is defined by (2.14). If such an ω exists, all other holomorphic differentials ω' are of the form $f\omega$ with $\operatorname{div}(f)\geq 0$, which implies $f\in K$; so, $g=\dim_K\Omega_{E/K}=1$, and $E_{/K}$ is an elliptic curve. It is an easy exercise to show that $y^{-1}dx$ does not vanish on E (Exercise 2).

We summarize what we have seen. Returning to the starting elliptic curve $E_{/K}$, for the parameter T at the origin, we see by definition

$$x=T^{-2}(1+\text{higher degree terms}) \quad\text{and}\quad y=-2T^{-3}(1+\text{higher degree terms}).$$

This shows

$$\frac{dx}{y}=\frac{-2T^{-3}(1+\cdots)}{-2T^{-3}(1+\cdots)}dT=(1+\text{higher degree terms})dT=\omega.$$

Thus the nowhere-vanishing differential form ω to which T is adapted is given by $\frac{dx}{y}$. Conversely, if $\Delta\neq 0$, the curve defined by $y^2=4x^3-g_2x-g_3$ is an elliptic curve over K with origin $\mathbf{0}=\infty$ and a standard nowhere-vanishing differential form $\omega=\frac{dx}{y}$. This finishes the proof.

Exercises

1. Show that dx/y does not vanish at any point on E.
2. Show that if $\Delta=0$, the curve defined by $y^2=4x^3-g_2x-g_3$ is not smooth at the multiple root α of $4x^3-g_2x-g_3=0$.

2.5.3 Moduli of Weierstrass Type

We continue to assume that the characteristic of K is different from 2 and 3. Suppose that we are given two elliptic curves $(E,\omega)_{/K}$ and $(E',\omega')_{/K}$ with nowhere-vanishing differential forms ω and ω'. We call two pairs (E,ω) and (E',ω') isomorphic if we have an isomorphism $\varphi : E \to E'$ with $\varphi^*\omega' = \omega$. Here for $\omega' = fdg$, $\varphi^*\omega' = (f\circ\varphi)d(g\circ\varphi)$; in other words, if $\sigma : \mathfrak{K}' \to \mathfrak{K}$ is the isomorphism of the function fields associated with φ, $\varphi^*\omega' = \sigma(f)d(\sigma(g))$. Let T' be the parameter at the origin $\mathbf{0}$ of E' adapted to ω'. If $\varphi : (E,\omega) \cong (E',\omega')$, then the parameter $T = \varphi^*T' \bmod T^2$ is adapted to ω (because $\varphi^*\omega' = \omega$). We choose coordinates (x,y) for E and (x',y') for E' relative to T and T' as above. By the uniqueness of the choice of (x,y) and (x',y'), we know $\varphi^*x' = x$ and $\varphi^*y' = y$. Thus the Weierstrass equations of (E,ω) and (E',ω') coincide. We write $g_2(E,\omega)$ and $g_3(E,\omega)$ for the g_2 and g_3 of the coefficients of the Weierstrass equation of (E,ω). Considering a polynomial ring $K[g_2,g_3]$ with variables g_2 and g_3, if K has a characteristic different from 2 and 3, we have

$$\left[(E,\omega)_{/K}\right] \cong \left\{(g_2,g_3)\in K^2 \middle| \Delta(E,\omega)\neq 0\right\} \cong Spec(\mathbb{Z}[X,Y,\frac{1}{X^3-27Y^2}])(K),$$

where $[\cdot]$ indicates the set of isomorphism classes of the objects inside the bracket and $Spec(R)(K)$ for a ring R is the set of all algebra homomorphisms: $R\to K$. The last isomorphism sends (g_2,g_3) to the algebra homomorphism ϕ with $\phi(X) = g_2$ and $\phi(Y) = g_3$.

We now classify elliptic curves E eliminating the contribution of the differential from the pair (E,ω). If $\varphi : E \cong E'$ for (E,ω) and (E',ω'), we have $\varphi^*\omega' = \lambda\omega$ with $\lambda\in K^\times$, because $\varphi^*\omega'$ is another nowhere-vanishing differential. Therefore we study $K^\times$-orbit: $(E,\omega) \bmod K^\times$ under the action of $\lambda\in K^\times$ given by $(E,\omega)_{/K} \longmapsto (E,\lambda\omega)_{/K}$, computing the dependence of $g_j(E,\lambda\omega)$ $(j=2,3)$ on λ for a given pair $(E,\omega)_{/K}$. Let T be the parameter adapted to ω. Then λT is adapted to $\lambda\omega$. We see

$$x(E,\omega) = \frac{(1+T\phi(T))}{T^2} \Rightarrow x(E,\lambda\omega) = \frac{(1+\text{higher terms})}{(\lambda T)^2} = \lambda^{-2}x(E,\omega),$$
$$y(E,\omega) = \frac{(-2+T\psi(T))}{T^3} \Rightarrow y(E,\lambda\omega) = \frac{(-2+\text{higher terms})}{(\lambda T)^3} = \lambda^{-3}y(E,\omega).$$

Since $y^2 = 4x^3 - g_2(E,\omega)x - g_3(E,\omega)$, we have

$$\begin{aligned}(\lambda^{-3}y)^2 &= 4\lambda^{-6}x^3 - g_2(E,\omega)\lambda^{-6}x - \lambda^{-6}g_3(E,\omega)\\ &= 4(\lambda^{-2}x)^3 - \lambda^{-4}g_2(E,\omega)(\lambda^{-2}x) - \lambda^{-6}g_3(E,\omega).\end{aligned}$$

This shows

$$g_2(E,\lambda\omega) = \lambda^{-4}g_2(E,\omega) \quad\text{and}\quad g_3(E,\lambda\omega) = \lambda^{-6}g_3(E,\omega). \tag{2.16}$$

Thus we have

Theorem 2.34 *If two elliptic curves $E_{/K}$ and $E'_{/K}$ are isomorphic, then choosing nowhere-vanishing differentials $\omega_{/E}$ and $\omega'_{/E'}$, we have $g_J(E', \omega') = \lambda^{-2J} g_J(E, \omega)$ for $\lambda \in K^\times$. The constant λ is given by $\varphi^*\omega' = \lambda\omega$.*

We define the J-invariant of E by $J(E) = \frac{(12g_2(E,\omega))^3}{\Delta(E,\omega)}$. Then J only depends on E (not the chosen differential ω). If $J(E) = J(E')$, then we have

$$\frac{(12g_2(E,\omega))^3}{\Delta(E,\omega)} = \frac{(12g_2(E',\omega'))^3}{\Delta(E',\omega')} \iff g_J(E',\omega') = \lambda^{-2J} g_J(E,\omega)$$

for a twelfth root λ of $\Delta(E,\omega)/\Delta(E',\omega')$. Note that the twelfth root λ may not be in K if K is not algebraically closed.

Conversely, for a given $j \notin \{0, 1\}$, the elliptic curve defined by $y^2 = 4x^3 - gx - g$ for $g = \frac{27j}{j-1}$ has J-invariant $12^3 j$. If $j = 0$ or 1, we can take the following elliptic curve with $J = 0$ or 12^3. If $J = 0$, then $y^2 = 4x^3 - 1$ and if $J = 12^3$, then $y^2 = 4x^3 - x$. Thus we have

Corollary 2.35 *If K is algebraically closed, then $J(E) = J(E') \Leftrightarrow E \cong E'$ for two elliptic curves over K. Moreover, for any field K, there exists an elliptic curve E with a given $J(E) \in K$.*

Exercises

1. Prove that $g_J(E', \omega') = \lambda^{-2J} g_J(E, \omega)$ for suitable ω and ω' and a suitable twelfth root λ of $\Delta(E,\omega)/\Delta(E',\omega')$ if $J(E) = J(E')$.
2. Explain what happens if $J(E) = J(E')$ but $E \ncong E'$ over a field K not necessarily algebraically closed.

2.5.4 Group Structure on Elliptic Curves

We now introduce algebrogeometric group structure on elliptic curves. First assume that K is algebraically closed. We know that $L(2[\mathbf{0}]) = K + Kx$. For a given point $P \in E - \{\mathbf{0}\}$, x is finite at P. We put $f = x - x(P)$. The function f has a zero at P. Since $\deg(\mathrm{div}(f)) = 0$ and f has at most an order 2 pole at $\mathbf{0}$, we have two possibilities: $\mathrm{div}(f) = P + P' - 2[\mathbf{0}]$ or $\mathrm{div}(f) = P - [\mathbf{0}]$. If the latter case happens, $[\mathfrak{K} : K(f)] = 1$ and hence $E \cong \mathbf{P}^1$, which is impossible since the genus of $\mathbf{P}^1$ is 0 (Exercise 1). Thus we find a unique P' from P, and we define $P' = -P$. When $P = \mathbf{0}$, the above argument just gives $P' = \mathbf{0}$; so, the definition of $P \mapsto -P$ is valid for all P.

Now we define addition of two points $P, Q \in E$. For $P, Q \in E - \{\mathbf{0}\}$, we solve a system of simultaneous linear equations:

$$\begin{cases} x(P)X + y(P)Y + Z = 0, \\ x(Q)X + y(Q)Y + Z = 0. \end{cases}$$

Here $(1, x, y)$ is a base of $L(3[\mathbf{0}])$. Then for a nontrivial solution (a, b, c) of the above system, we define $g = ax + by + c$. We assume that $Q \neq -P$. Then $b \neq 0$ (because $b = 0 \Rightarrow Q = -P$ as is clear from the definition of $-P$). Thus g has a pole of order 3 at $\mathbf{0}$. Since g is nonconstant with two zeros P, Q, we find a unique point R such that $\mathrm{div}(g) = P + Q + R - 3[\mathbf{0}]$. Then we define $P + Q$ to be $-R$. By Abel's theorem we prove in the following subsection, this addition gives rise to an abelian group structure on E. When $Q = -P$, we just define $P + Q = \mathbf{0}$.

Here is a geometric interpretation of the above definition of the addition. Embed E into $\mathbf{P}^2$ by $1, x, y$. Then we draw a line $aX + bY + cZ = 0$ passing through P, Q in $\mathbf{P}^2$, where (X, Y, Z) is the homogeneous coordinate of $\mathbf{P}^2$. Then we consider the function $\phi : E \to \mathbf{P}^1$ given by the equation $\phi(X, Y, Z) = (aX + bY + cZ)/Z$, which is an element in $\mathfrak{K} = K(E)$. We see that ϕ vanishes at two points P, Q, and the line intersects with E at a unique point R again. Then $\mathrm{div}(\phi) = P + Q + R - 3[\mathbf{0}]$. In particular, the coordinate of R is a rational function of the coordinates of P and Q; so, $+ : E \times E \to E$ is a morphism of algebraic geometry (in other words, it is induced from a field embedding $\mathfrak{K} \hookrightarrow \mathfrak{K} \otimes_K \mathfrak{K}$). Similarly, $P \mapsto -P$ is an automorphism of the elliptic curve (therefore of the algebraic function field $\mathfrak{K}$).

After embedding E into $\mathbf{P}^2$, for any field automorphism σ of K, we can apply σ to the coordinate of E. Then we get a new elliptic curve E^σ. If E is defined, for example, by $F(X, Y) = \sum_{i,j} a_{i,j} X^i Y^j = 0$, then E^σ is defined by $F^\sigma(X, Y) = \sum_{i,j} \sigma(a_{i,j}) X^i Y^j = 0$. Since everything we have proved for E shifts to E^σ, the morphism $+ : E \times E \to E$ will be sent to $+ : E^\sigma \times E^\sigma \to E^\sigma$. If E is defined over a smaller field $k \subset \overline{k} = K$, the rational functions of the coordinates of E giving rise to $+ : E \times E \to E$ are therefore invariant under σ; so, it is a rational function with coefficients in k. Thus what we have said so far is valid for any elliptic curve defined over any perfect field (not necessarily over an algebraically closed field).

Exercises

1. Prove that $\mathbf{P}^1$ has genus 0.
2. Give a detailed argument why the morphism giving addition on E is well-defined over its field of definition.

2.5.5 Abel's Theorem

We now prove the following Abel's theorem for elliptic curves.

Theorem 2.36 *Let E be an elliptic curve with origin $\mathbf{0}$ defined over a perfect field K. Then $i : E(K) \to \mathrm{Jac}(E)(K)$ given by $i(P) = P - [\mathbf{0}]$ is an isomorphism of groups.*

By this theorem, we find that $P + Q + R \sim 3[\mathbf{0}]$ (here $\sim$ is the linear equivalence) if and only if $i(P) + i(Q) + i(R) = 0$ in $\mathrm{Jac}(E)$; so, the group structure

we have defined in the previous subsection is induced from the group structure of the Jacobian $\mathrm{Jac}(E)$.

Proof. By the same argument (as in the last part of the previous subsection) conjugating by automorphisms of K, everything we prove over an algebraically closed field K will be valid over any subfield k such that K/k is algebraic. Thus we may assume that K is algebraically closed. We first remark that

$$L(D) \neq \{0\} \Rightarrow \deg(D) \geq 0. \tag{2.17}$$

Indeed, if $0 \neq f \in L(D)$ implies that $\mathrm{div}(f) + D \geq 0$; so, taking the degree, we have $\deg(D) \geq -\deg(\mathrm{div}(f)) = 0$. This shows

$$\deg(D) < 0 \Rightarrow \dim_K L(D) = 0. \tag{2.18}$$

Take a degree 1 divisor $D = \sum_{i=1}^{d} P_i - \sum_{j=1}^{d-1} Q_j$. By the Riemann–Roch theorem, we have $\dim L(D) = \deg(D) + \dim L(-D) = \deg(D) = 1$ by (2.18). We find a nonconstant $f \in L(D)$. This f has to vanish at all Q_j. We have two possibilities:

- $\mathrm{div}(f) = \sum_{j=1}^{d-1} Q_j - \sum_{i=1}^{d-1} P_i$ after renumbering P_i; so, $D \sim P_d$;
- f has an extra zero R. Then $\mathrm{div}(f) = \sum_{j=1}^{d-1} Q_j + R - \sum_{i=1}^{d} P_i$; so, $D \sim R$.

Thus D is linearly equivalent to one point. In other words, for any divisor D_0 of degree 0, $D_0 + [\mathbf{0}] \sim P$ for a point P, and hence, the class of D_0 is in the image of i; so, i is surjective. If $i(P) = i(Q)$ for $P \neq Q$, then we have $P - Q \sim 0$, and hence $\dim L(P - Q) = \dim L(0) = 1$. Therefore we have a function f with a simple pole at P and a simple zero at Q. Thus $[\mathfrak{K} : K(f)] = 1$ and $E \cong \mathbf{P}^1$, a contradiction; so, i is injective. □

2.5.6 Torsion Points on Elliptic Curves

Let T be a parameter of E over K at the origin $\mathbf{0}$. Then taking a copy S of T, we may think that S, T is a set of parameters of $E \times E$ at $(\mathbf{0}, \mathbf{0})$. Then $(P, Q) \mapsto f(P + Q)$ gives rise to a meromorphic function on $E \times E$. In other words, the function is a pullback of f by $+$. In particular, we have $T \circ + = \Phi \in K[[T, S]]$. For $f(T)$, we find $f \circ +(T, S) = f(\Phi(T, S))$, and by associativity, we get $\Phi(T, \Phi(S, U)) = \Phi(\Phi(T, S), U)$ (Exercise 1). In particular, sending $(T, S) \mapsto \Phi(T, S) = aT + bS \in (T, S)/(T, S)^2$, we get a new associative and commutative group law on K, which is linear with respect to K. Since such a group law is unique on the field K, we find that $\Phi(T, S) \equiv T + S \mod (T, S)^2$ and, hence, the addition of the elliptic curve coincides with the usual addition on K in this sense.

Let N be a positive integer. Then $x \mapsto N \cdot x = \overbrace{x + x + \cdots + x}^{N}$ is a morphism of E into itself, and N induces on the vector space $(T)/(T^2)$ the multiplication by N. The morphism $x \mapsto N \cdot x$ induces an embedding N^* of the

function field $\mathfrak{K} = K(E)$ into itself. The degree $[\mathfrak{K} : N^*\mathfrak{K}] = \deg(N)$ is finite, because $\mathfrak{K}$ is finitely generated over K and of transcendental degree 1. Thus we find that $N : E \to E$ is surjective for all positive integer N. In particular, $E(K)$ is an N-divisible group if K is algebraically closed, and we have proven this fact as long as N is prime to the characteristic of K.

A naive question is: What is $\deg(N) = [\mathfrak{K} : N^*(\mathfrak{K})]$? The following theorem gives us that information and slightly more.

Theorem 2.37 *Suppose that K is algebraically closed. Then $\deg(N) = N^2$. Moreover, if N is outside the characteristic of K, we have $E[N](K) \cong (\mathbb{Z}/N\mathbb{Z})^2$ as abelian groups, and if $p > 0$ is the characteristic of the field K, then $E[p]$ is either trivial or isomorphic to $\mathbb{Z}/p\mathbb{Z}$.*

We prove this theorem in the following subsection when K is a subfield of $\mathbb{C}$. See [GME] 2.6 for more general cases.

In the rest of this subsection, we explore consequences of the theorem. Start with an arbitrary field K and write $\overline{K}$ for its algebraic closure. Take an elliptic curve E defined over K with function field $\mathfrak{K}$. We write the composite $\mathfrak{K}\overline{K}$ as $\overline{\mathfrak{K}}$, which gives rise to an elliptic curve $\overline{E}$ over $\overline{K}$. As the algebraic curve inside $\mathbf{P}^m$ (the integer m can be taken to be 2 as we have seen), E is just defined by equations giving relations among generators of $\mathfrak{K}$ over K. Since $\overline{\mathfrak{K}}$ has the same generators (as $\mathfrak{K}$) over $\overline{K}$, they give rise to the same algebraic curve inside $\mathbf{P}^m$. For any extension F/K inside $\overline{K}$, we define $E(F)$ to be the set of points of $\overline{E}$ with coordinates in F. Thus $E(F)$ is the subset of $\overline{E}(\overline{K})$ fixed by $\mathrm{Gal}(\overline{K}/F)$.

Take two points P, Q in $E(\overline{K})$. Since $P + Q$ is determined by the third intersection with E of the line passing through P and Q, if we conjugate the coordinate by $\sigma \in \mathrm{Gal}(\overline{K}/K)$, the resulting image goes to the third intersection with E of the line passing through $\sigma(P)$ and $\sigma(Q)$ (when $P = Q$, the line is the line tangent to $P \in E(\overline{K})$). This shows that $\sigma(P) + \sigma(Q) = \sigma(P + Q)$, and conjugation by Galois automorphism is compatible with the addition of E. In particular, $\sigma(N \cdot P) = N \cdot \sigma(P)$, so the Galois action preserves $E[N]$ (since $\sigma(\mathbf{0}) = \mathbf{0}$ because $\mathbf{0} \in E(K)$). In other words, $\sigma : E[N] \to E[N]$ is an injective group homomorphism. Since $E[N]$ is a finite group, σ induces an automorphism of the group $E[N]$.

Let P, Q be a base of $E[N]$ over $\mathbb{Z}/N\mathbb{Z}$. Then $\sigma(P), \sigma(Q)$ is another base. Write $\rho(\sigma)$ the base-change matrix:

$$\begin{pmatrix} \sigma(P) \\ \sigma(Q) \end{pmatrix} = \begin{pmatrix} aP+bQ \\ cP+dQ \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} P \\ Q \end{pmatrix}$$

with $\rho(\sigma) = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. By definition, we find $\rho(\sigma)\rho(\tau) = \rho(\sigma\tau)$ and $\rho(1) = 1_2$, where $1_2 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$.

Theorem 2.38 *Let E be an elliptic curve defined over K. Suppose that the integer N is outside the characteristic of K. Then $\sigma \mapsto \rho(\sigma)$ is a homomorphism $\rho : \mathrm{Gal}(\overline{K}/K) \to GL_2(\mathbb{Z}/N\mathbb{Z})$, where $GL_2(\mathbb{Z}/N\mathbb{Z})$ is the group*

of invertible two-by-two matrices with coefficients in the ring $\mathbb{Z}/N\mathbb{Z}$. *Moreover,* $\det\rho(\sigma) = \chi_N(\sigma)$, *where* $\chi_N(\sigma)$ *is determined by* $\sigma(\zeta_N) = \zeta_N^{\chi_N(\sigma)}$ *for a primitive root of unity* ζ_N.

By the above theorem, the field $F = K(E[N])$ generated by coordinates of all points of $E[N]$ is a Galois extension of K with $\mathrm{Gal}(F/K) \cong \mathrm{Im}(\rho) \subset GL_2(\mathbb{Z}/N\mathbb{Z})$, and F contains all Nth roots of unity. We later explore some cases where ρ is surjective.

Proof. We only need to prove the last assertion about $\det\rho = \chi_N$ for the cyclotomic character χ_N. We construct a bilinear form $e = e_N : E[N] \times E[N] \to \mu_N$ for the group of Nth roots of unity μ_N such that

(P1) $e(x,x) = 1$ for all $x \in E[N]$ (alternating);
(P2) $e(x, E[N]) = 1 \Rightarrow x = \mathbf{0}$ (nondegeneracy);
(P3) $\sigma(e(x,y)) = e(\sigma(x), \sigma(y))$ (Galois equivariance).

If this is worked out, we find $e(P,Q) = \zeta_N$ is a primitive Nth root (P2) and

$$\begin{aligned}\sigma(\zeta_N) = \sigma(e(P,Q)) &= e(\sigma(P), \sigma(Q)) \\ &= e(aP+bQ, cP+dQ) = e(P,Q)^{ad} e(Q,P)^{bc} = \zeta_N^{ad-bc}\end{aligned}$$

as desired (because $1 = e(P+Q, P+Q) = e(P,Q)e(Q,P)$). We need to construct e and to prove (P1–3) above. Here is a construction given by A. Weil: we need to consider a point $P \in E$ both as a geometric point and also a divisor; so, we write $[P]$ when we consider P as a divisor. By Abel's theorem,

(div) if $\sum_i c_i x_i = \mathbf{0}$ for $x_i \in E(\overline{K})$ and $\sum_i c_i = 0$ for integers c_i, then $0 = \sum_i c_i([x_i] - [\mathbf{0}]) = \sum_i c_i [x_i]$.

Pick $x \in E[N]$; so, $N[x] - N[\mathbf{0}] = \mathrm{div}(f_x)$ for a function $f_x \in \overline{\mathfrak{K}}$. Since $E(\overline{K})$ is divisible, we can pick $t \in E(\overline{K})$ such that $x = Nt$. We consider the divisor $D = \sum_{u \in E[N]} [t+u] - \sum_{u \in E[N]} [u]$. Then by (div), we find a function $g_x \in \overline{\mathfrak{K}}$ such that $\mathrm{div}(g_x) = D$. We see that $\mathrm{div}(f_x \circ N)$ is made up of points that go to $\mathbf{0}$ and t with multiplicity N; so, $\mathrm{div}(f_x \circ N) = N \cdot \mathrm{div}(g_x)$. Thus $g_x^N / f_x \circ N$ is a nonzero constant c, and $g_x(P+y)^N = c f_x(N(P+y)) = c f_x(NP) = g_x(P)^N$ for $y \in E[N]$. We then define $e(x,y) = g_x(P+y)/g_x(P)$ which is an Nth roots of unity. By this construction, $g_{\sigma(x)} = c_1 \sigma(g_x)$ for a nonzero constant $c_1 \in \overline{K}$; so, we may choose $g_{\sigma(x)}$ to be $\sigma(g_x)$. Then Galois equivariance follows from

$$\begin{aligned}\sigma(e(x,y))\sigma(g_x(P)) &= \sigma(g_x(P+y)) \\ &= \sigma(g_x)(\sigma(P) + \sigma(y)) = e(\sigma(x), \sigma(y))\sigma(g_x(P)).\end{aligned}$$

For $y, y' \in E[N]$, $g_x(P+y+y') = e(x,y')g_x(P+y) = e(x,y)e(x,y')g_x(P)$; so, the pairing is linear with respect to the right variable. We now take care of the linearity for the left variable. Let $z = x + w$ for $x, w \in E[N]$. Since $x + w - z - \mathbf{0} = \mathbf{0}$, applying (div), we find $h \in \overline{\mathfrak{K}}$ such that $\mathrm{div}(h) = [x] + [w] - [z] - [\mathbf{0}]$. Then

$$\mathrm{div}(f_x f_w f_z^{-1}) = N([x] - [\mathbf{0}] + [w] - [\mathbf{0}] + [\mathbf{0}] - [z]) = N \cdot \mathrm{div}(h)$$

and $N \cdot \mathrm{div}(g_x g_w g_z^{-1}) = N \cdot \mathrm{div}(h \circ N)$. Thus $g_x g_w g_z^{-1} = c(h \circ N)$ for a nonzero constant $c \in K$. In other words,

$$\begin{aligned} &e(x,y)^{-1} e(w,y)^{-1} e(z,y) g_x(P) g_w(P) g_z^{-1}(P) \\ &\quad = g_x(P+y) g_w(P+y) g_z^{-1}(P+y) = c \cdot h(P) = g_x(P) g_w(P) g_z^{-1}(P), \end{aligned}$$

which shows the desired left linearity.

If $e(x,y) = 1$ for all $y \in E[N]$, g_x factors through the image of N; that is, $g_x = g' \circ N$, and hence $f_x = {g'}^N$ (so, $N \cdot \mathrm{div}(g') = \mathrm{div}(f_x)$) because $N : E \to E$ is surjective. We have $\mathrm{div}(g') = [x] - [\mathbf{0}]$, and by Abel's theorem $x = \mathbf{0}$. This proves the nondegeneracy.

We now prove (P1). Observe

$$\mathrm{div}\Big(\prod_{n=0}^{N-1} f_x(P - nx)\Big) = N \sum_{n=0}^{N-1} ([nx + x] - [nx]) = 0.$$

Then $\prod_{n=0}^{N-1} f_x(P - nx)$ is a constant. Since we have chosen t so that $Nt = x$, we have $g_x(P - nt) = f_x(N(P - nt))^N = f_x(NP - nx)^N$, and $\prod_{n=0}^{N-1} g_x(P - nt)$ has to be a constant. Now we plug $P - x$ in P, and we get

$$\begin{aligned} &g_x(P) g_x(P-t) \cdots g_x(P-(N-1)t) \\ &\qquad = g_x(P-t) g_x(P-2t) \cdots g_x(P-(N-1)t) g_x(P-x). \end{aligned}$$

This shows that $g_x(P) = g_x(P - x)$ and hence $e(x,x) = 1$. □

Exercises

1. Prove that $\Phi(T, \Phi(S,R)) = \Phi(\Phi(T,S), R)$ in the power series ring of three variables T, S, R over K.
2. Prove $\rho(\sigma)\rho(\tau) = \rho(\sigma\tau)$.

2.5.7 Classical Weierstrass Theory

Let $E_{/\mathbb{C}}$ be an elliptic curve over $\mathbb{C}$. Since E is of genus 1, it is a quotient of $\mathbb{C}$ by a lattice $L \subset \mathbb{C}$. Here a lattice $L \subset \mathbb{C}$ is a $\mathbb{Z}$-submodule of $\mathbb{C}$ generated by a base (w_1, w_2) of $\mathbb{C}$ over $\mathbb{R}$. We write Lat for the set of all lattices in $\mathbb{C}$.

Taking a nowhere-vanishing differential ω on $E(\mathbb{C})$, we rediscover the lattice L of $\mathbb{C}$ (associated with (E, ω)) by

$$L = \left\{ \int_\gamma \omega \middle| \gamma \in H_1(E(\mathbb{C}), \mathbb{Z}) \right\}.$$

Abel's isomorphism is then given by $P \mapsto \int_0^P \omega \in \mathbb{C}/L = \mathrm{Jac}(E)$. By Abel's theorem, we find that $E[N] = \frac{1}{N}L/L \cong (\mathbb{Z}/N\mathbb{Z})^2$, which gives a proof of Theorem 2.37 if $K = \mathbb{C}$. For a general field K of characteristic 0 inside $\mathbb{C}$ and an elliptic curve $E_{/K}$, the composite $\mathbb{C}\mathfrak{K}$ gives rise to the elliptic curve $E_{\mathbb{C}}$ over $\mathbb{C}$ defined by the same equation; so, $E[N](\overline{K}) \cong E_{\mathbb{C}}[N]$, which proves the theorem basically for all fields of characteristic 0.

Conversely, for a given $L \in Lat$, we define the Weierstrass $\mathcal{P}$-functions by

$$x_L(u) = \mathcal{P}(u) = \frac{1}{u^2} + \sum_{\ell \in L - \{0\}} \left\{ \frac{1}{(u-\ell)^2} - \frac{1}{\ell^2} \right\} = \frac{1}{u^2} + \frac{g_2}{20}u^2 + \frac{g_3}{28}u^4 + \cdots,$$

$$y_L(u) = \mathcal{P}'(u) = \frac{-2}{u^3} - 2 \sum_{\ell \in L - \{0\}} \frac{1}{(u-\ell)^3} = -2u^{-3} + \cdots,$$

where

$$g_2 = g_2(L) = 60 \sum_{\ell \in L - \{0\}} \frac{1}{\ell^4} \quad \text{and} \quad g_3 = g_3(L) = 140 \sum_{\ell \in L - \{0\}} \frac{1}{\ell^6}.$$

Then $\varphi = y_L^2 - 4x_L^3 + g_2 x_L + g_3$ is holomorphic everywhere. Since these functions factor through the compact space $\mathbb{C}/L$, φ has to be constant, because any nonconstant holomorphic function is an open map (the existence of power series expansion and the implicit function theorem). Since x_L and y_L do not have constant terms, we conclude $\varphi = 0$. We have obtained a holomorphic map $(x_L, y_L) : \mathbb{C}/L - \{0\} \to \mathbb{C}^2_{/\mathbb{C}}$. Looking at the order of poles at $\mathbf{0}$, we know the above map is of degree 1, that is, an isomorphism onto its image. It also extends to the embedding $\Phi = (x_L, y_L, 1) = (u^3 x_L, u^3 y_L, u^3) : \mathbb{C}/L \to \mathbf{P}^2_{/\mathbb{C}}$. Thus we get an elliptic curve $E_L = \Phi(\mathbb{C}/L) = E(g_2(L), g_3(L))$. We then have

$$\omega_L = \frac{dx_L}{y_L} = dx_L \left(\frac{dx_L}{du} \right)^{-1} = du.$$

This shows

Theorem 2.39 (Weierstrass) *We have*

$$Lat \cong \left[(E, \omega)_{/\mathbb{C}} \middle| \begin{array}{l} E\text{: an elliptic curve over } \mathbb{C}, \\ \omega\text{: a nowhere-vanishing differential} \end{array} \right],$$

where the straight brackets [ ] *indicate the set of all isomorphism classes of the objects inside.*

We now make the space *Lat* more explicit. Two complex numbers w_1, w_2 span a lattice if and only if $\mathrm{Im}(w_1/w_2) \neq 0$. Let $\mathfrak{H} = \{z \in \mathbb{C} | \, \mathrm{Im}(z) > 0\}$. Interchanging w_1 and w_2 if necessary, we may assume that $\mathrm{Im}(w_1/w_2) > 0$. So we get a natural isomorphism of complex manifolds via $\left(\begin{smallmatrix} w_1 \\ w_2 \end{smallmatrix}\right) \mapsto (w_2, w_1/w_2)$:

$$\mathcal{B} = \left\{ v = \left(\begin{smallmatrix} w_1 \\ w_2 \end{smallmatrix}\right) \in (\mathbb{C}^\times)^2 \,\middle|\, \mathrm{Im}(w_1/w_2) > 0 \right\} \cong \mathbb{C}^\times \times \mathfrak{H}.$$

Since v and v' span the same lattice L if and only if $v' = \alpha v$ for $\alpha \in SL_2(\mathbb{Z})$, we see that $Lat \cong SL_2(\mathbb{Z})\backslash\mathcal{B}$. The action of $\alpha = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z})$ on $\mathcal{B}$ is given on $\mathbb{C}^\times \times \mathfrak{H}$ by $\alpha(u, z) = (cu + d, \alpha(z))$ for $\alpha(z) = \frac{az+b}{cz+d}$.

We now consider the variables g_2 and g_3 and the J-invariant J defined below Theorem 2.34 as a function on the upper half complex plane $\mathfrak{H}$. In particular, by Corollary 2.35, J satisfies $J(\gamma(z)) = J(z)$ for $\gamma \in SL_2(\mathbb{Z})$.

2.6 Elliptic Modular Function Field

Combining the analytic consideration in the previous subsection with the algebraic study in the earlier ones, we now state and partially prove the first non-abelian global reciprocity law due to Shimura.

For a lattice L spanned by (w_1, w_2) with $\mathrm{Im}(w_1/w_2) > 0$, let $E(L)$ be the elliptic curve given by $E(\mathbb{C}) = \mathbb{C}/L$. Picking a point $u(a, b) = \frac{aw_1+bw_2}{N} \in E(L)[N]$ (for $0 \neq u = (a, b) \in (\mathbb{Z}/N\mathbb{Z})^2$), we define a function $f_u : Lat \to \mathbb{C}$ by

$$f_u(w) = \frac{g_2(w)g_3(w)}{\Delta(w)} x_L(u(a, b)) \text{ and } x_u(w) = x_L(u(a, b))$$

for $w = {}^t(w_1, w_2)$ and $\Delta = g_2^3 - 27g_3^2$. Then for $\alpha \in SL_2(\mathbb{Z})$,

$$f_u(\alpha w) = f_u(w) \text{ for } 0 \neq \forall u = (a, b) \in (\mathbb{Z}/N\mathbb{Z})^2 \Leftrightarrow \alpha \equiv \pm 1_2 \mod N. \tag{2.19}$$

Also we see easily that $x_u(\lambda w) = \lambda^{-2} x_u(w)$ for nonzero scalar $\lambda \in \mathbb{C}^\times$. We call a meromorphic function f on $\mathcal{B}$ a modular form of weight k on $\Gamma(N)$ if f satisfies $f(\gamma w) = f(w)$ for all $\gamma \in SL_2(\mathbb{Z})$ with $\gamma - 1_2 \in NM_2(\mathbb{Z})$ and $f(\lambda w) = \lambda^{-k} f(w)$ for $\lambda \in \mathbb{C}^\times$. In particular, if $k = 0$, they are called modular functions on $\Gamma(N)$. Thus f_u is a modular function on $\Gamma(N)$, and x_u is a modular form of weight 2 on $\Gamma(N)$. We often let $\Gamma(N)$ denote the subgroup of $SL_2(\mathbb{Z})$ made of matrices congruent to 1_2 modulo N. We have the following exact sequence,

$$1 \to \Gamma(N) \to SL_2(\mathbb{Z}) \xrightarrow{\mod N} SL_2(\mathbb{Z}/N\mathbb{Z}) \to 1.$$

The surjectivity of the last "mod N" map is nontrivial (Exercise 1).

Since $\mathcal{B} = \mathfrak{H} \times \mathbb{C}^\times$, one may regard a modular form as a meromorphic function on the upper half complex plane $\mathfrak{H}$ by putting $f(z) = f\left(2\pi i \left(\begin{smallmatrix} z \\ 1 \end{smallmatrix}\right)\right)$. Since $\alpha \left(\begin{smallmatrix} z \\ 1 \end{smallmatrix}\right) = (cz + d) \left(\begin{smallmatrix} \alpha(z) \\ 1 \end{smallmatrix}\right)$ for $\alpha = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, we find that $f : \mathfrak{H} \to \mathbb{C}$ is a modular form of weight k on $\Gamma(N)$ if the following conditions are satisfied,

(G1) $\quad f(\alpha(z)) = f(z)(cz + d)^k$ for all $\alpha = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma(N)$.

If f is a modular form on $\Gamma(N)$, $f(z+N) = f(z)$; so, f is a function on $\mathfrak{H}/N\mathbb{Z}$, where $N\mathbb{Z}$ acts by translation. By the variable change, $q = q_N = \exp(\frac{2\pi\imath z}{N})$,

we may identify $\mathfrak{H}/N\mathbb{Z}$ with the open unit disk $\{q \in \mathbb{C} | |q| < 1\}$ punctured at 0 (0 corresponds to ∞), and we may regard f as a function of q defined over the punctured disk. It has a Laurent expansion around 0: $f(q) = \sum_n a_n q^n$, which is called the *q-expansion* of f. We assume that $f(q)$ is finite tailed; that is, there exists $N > 0$ such that $a_n = 0$ as long as $n < -N$.

We can compute explicitly the q-expansion of g_2, g_3 and Δ for $q = q_1$:

$$\begin{aligned} 12g_2 &= 1 + 240 \sum_{n=1}^{\infty} \left\{ \sum_{0<d|n} d^3 \right\} q^n \\ -6^3 g_3 &= 1 - 504 \sum_{n=1}^{\infty} \left\{ \sum_{0<d|n} d^5 \right\} q^n \\ \Delta &= q \prod_{n=1}^{\infty} (1 - q^n)^{24}. \end{aligned} \tag{2.20}$$

From this we can conclude $J \in q^{-1}\mathbb{Z}[[q]]$ (cf. [IAT] Section 4.6).

We consider the Riemann surface $Y(N) = \Gamma(N)\backslash\mathfrak{H}$. Regarding $J = 12^3 j$ as a variable, and considering the elliptic curve $E : y^2 = 4x^3 - gx - g$ for $g = \frac{27j}{j-1}$ defined over $\mathbb{Q}(J)$, we can think of $\mathbb{Q}(J)(E[N])$. This is a Galois extension of $\mathbb{Q}(J)$ with

$$\mathrm{Gal}(\mathbb{Q}(J)(E[N])/\mathbb{Q}(J)) \subset GL_2(\mathbb{Z}/N\mathbb{Z}).$$

We also consider the field $\mathfrak{K}_N$ generated over $\mathbb{Q}(J)$ (now J is a function on $\mathfrak{H}$) by f_u for all $u \in (\mathbb{Z}/N\mathbb{Z})^2$. Since f_u is the x-coordinate of the point corresponding to $u = u(a, b) \in E[N]$ (up to the factor $g^2/g^3 - 27g^2$ in $\mathbb{Q}(J)$), we may regard $\mathfrak{K}_N \subset \mathbb{Q}(J)(E[N])$ and $\mathfrak{K}_N \subset \mathbb{C}(Y(N))$. It is easy to see that $\alpha \in SL_2(\mathbb{Z})$ acts on $Y(N)$ by $z \mapsto \alpha(z)$; so, it induces an automorphism of $\mathbb{C}(Y(N))/\mathbb{C}(J)$.

Theorem 2.40 *We have* $\mathrm{Gal}(\mathfrak{K}_N/\mathbb{Q}(J)) \cong GL_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$, *and* $\alpha \in SL_2(\mathbb{Z})$ *acts on* $f \in \mathfrak{K}_N$ *by* $f \mapsto f \circ \alpha$. *This action of* $SL_2(\mathbb{Z})$ *factors through* $PSL_2(\mathbb{Z}/N\mathbb{Z})$, *which is the subgroup of* $\mathrm{Gal}(\mathfrak{K}_N/\mathbb{Q}(J))$.

Proof. We check $f_u \circ \alpha = f_{u\alpha}$. Since $f_{u\alpha} = f_u$ for all u implies that $u\alpha = \pm u$ by definition, we find that $\mathrm{Gal}(\mathbb{C}(Y(N))/\mathbb{C}(J))$ contains $PSL_2(\mathbb{Z}/N\mathbb{Z})$. Since $Y(N)$ is the covering of degree $|PSL_2(\mathbb{Z}/N\mathbb{Z})|$ of $\mathbf{P}^1(J)$, we find that

$$\mathbb{C}(Y(N)) = \mathbb{C}(J)(f_u | 0 \neq u \in (\mathbb{Z}/N\mathbb{Z})^2).$$

By computing the q-expansion of f_u, we find that $f_u \in \mathbb{Q}[\zeta_N]((q))$:

$$x_u = \frac{1}{12} - 2\sum_{n=1}^{\infty} \frac{nq^{Nn}}{1 - q^{Nn}} + \frac{\zeta_N^b q^a}{(1 - \zeta_N^b q_N^a)^2} + \sum_{n=1}^{\infty} \frac{(\zeta_N^{bn} q^{an} + \zeta_N^{-bn} q^{-an}) n q^n}{1 - q^n}$$

for $u = (a, b)$ and $q = \exp(2\pi iz/N)$. By Theorem 2.38, we know that $\mathfrak{K}_N \supset \mathbb{Q}[\zeta_N]$; so, $\mathbb{C} \cap \mathfrak{K}_N = \mathbb{C} \cap \mathbb{Q}[\zeta_N]((q)) = \mathbb{Q}[\zeta_N]$. Since $\mathrm{Gal}(\mathfrak{K}_N/\mathbb{Q}(J))$ contains all $PSL_2(\mathbb{Z}/N\mathbb{Z})$ and also matrices with any given determinant modulo N, we find

$$\mathrm{Gal}(\mathfrak{K}_N/\mathbb{Q}(J)) = GL_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}.$$

This finishes the proof.

By the above proof, q gives a parameter at ∞ of $Y(N)$ defined over $\mathbb{Q}(\zeta_N)$; so, we have

$$\mathfrak{K}_N = \mathbb{C}(Y(N)) \cap \mathbb{Q}[\zeta_N]((q)) \quad \left(q = \exp\left(\frac{2\pi iz}{N}\right)\right). \tag{2.21}$$

This shows

Corollary 2.41 *The curve $Y(N)_{/\mathbb{C}}$ is actually defined over $\mathbb{Q}[\zeta_N]$, and*

$$\mathbb{Q}[\zeta_N](Y(N)_{/\mathbb{Q}[\zeta_N]}) = \mathfrak{K}_N.$$

We consider the union $\mathfrak{K} = \bigcup_N \mathfrak{K}_N$ (note here that $\mathfrak{K}_N \subset \mathfrak{K}_M$ if $N|M$); so, $\mathfrak{K}$ is a field.

Corollary 2.42 *We have $\mathrm{Gal}(\mathfrak{K}/\mathbb{Q}(J)) \cong GL_2(\widehat{\mathbb{Z}})/\{\pm 1\}$ and $\mathfrak{K} \cap \mathbb{C} = \mathbb{Q}^{cyc}$. Moreover, each matrix $g \in GL_2(\widehat{\mathbb{Z}})$ acts on $\mathbb{Q}^{cyc}$ by the action of $\det(g)$ under the identification $\mathrm{Gal}(\mathbb{Q}^{cyc}/\mathbb{Q}) = \widehat{\mathbb{Z}}^\times$.*

For $\alpha \in M_2(\mathbb{Z})$ with $\det(\alpha) > 0$, we can always write by elementary divisor theory, $\alpha = \gamma \left(\begin{smallmatrix} m & 0 \\ 0 & n \end{smallmatrix}\right) \gamma'$ with $\gamma, \gamma' \in SL_2(\mathbb{Z})$. We write $\delta = \left(\begin{smallmatrix} m & 0 \\ 0 & n \end{smallmatrix}\right)$. Since the effect of $f \mapsto f \circ \delta$ on q-expansion is $q \mapsto q^{m/n}$, the action $f \mapsto f \circ \alpha$ preserves the coefficient field $\mathbb{Q}[\zeta_N]$. Since $\delta^{-1}\Gamma(N)\delta \supset \Gamma(mnN)$ (Exercise 2), $\mathfrak{K}_N \circ \delta \subset \mathfrak{K}_{mnN}$. Since $\gamma, \gamma' \in SL_2(\mathbb{Z})$ preserve $\mathfrak{K}_N$, we see that $\mathfrak{K}_N \circ \alpha \subset \mathfrak{K}_{mnN}$ and hence $f \mapsto f \circ \alpha$ is an automorphism $\tau(\alpha)$ of $\mathfrak{K}$. For any $\alpha \in GL_2(\mathbb{Q})_+$, its integer multiple is in $M_2(\mathbb{Z})$. Thus the linear transformation $z \mapsto \alpha(z)$ is induced by an integer matrix; so, $\tau(\alpha) \in \mathrm{Aut}(\mathfrak{K})$ is well defined on $\mathfrak{K}$; that is, $GL_2(\mathbb{Q})_+$ acts on $\mathfrak{K}$ by $f \mapsto f \circ \alpha$. We can check (e.g., [IAT] 6.4 or [MFG] Corollary 3.3):

(A1) $GL_2(\mathbb{Q})_+ GL_2(\widehat{\mathbb{Z}}) GL_2(\mathbb{R})_+ = GL_2(\mathbb{A})_+$, where $+$ indicates the positivity of the determinant (at ∞);

(A2) $GL_2(\widehat{\mathbb{Z}} \times \mathbb{R})_+ \cap \mathbb{Q}^\times = \{\pm 1\}$.

Thus we can let $GL_2(\mathbb{A})_+$ act on $\mathfrak{K}$. Here is a theorem of Shimura.

Theorem 2.43 (Elliptic Reciprocity Law) *We have an exact sequence*

$$1 \to \mathbb{Q}^\times GL_2(\mathbb{R})_+ \to GL_2(\mathbb{A})_+ \xrightarrow{\tau} \mathrm{Aut}(\mathfrak{K}) \to 1.$$

The action of $GL_2(\mathbb{A})_+$ on $\mathfrak{K}$ is as given above. The difficult part of the proof is the surjectivity of τ for which we refer the reader to [IAT] Theorem 6.23. We later give a sketch of a proof in a more general setting, Theorem 4.14.

For each open compact subgroup S of $GL_2(\mathbb{A})_+/\mathbb{Q}^\times GL_2(\mathbb{R})_+$, we have an algebraic function field $\mathfrak{K}^S$, which gives rise to a projective curve V_S defined over the fixed subfield k_S of $\mathbb{Q}^{cyc}$ by $\det(S) \subset \mathbb{A}^\times/\mathbb{Q}^\times\mathbb{R}_+^\times = \mathrm{Gal}(\mathbb{Q}^{cyc}/\mathbb{Q})$. This curve is the modular curve of level S, and $V_S(\mathbb{C}) = \Gamma_S\backslash\mathfrak{H}$ (forgetting the points above $\infty \in \mathbf{P}^1(J)$). Here $\Gamma_S = S \cap SL_2(\mathbb{Q})$. In particular, $Y(N)$ corresponds to $\{u \in GL_2(\widehat{\mathbb{Z}}) | u \equiv 1_2 \mod N\}$, and we have the following identity,

$$\left[(E, \phi : (\mathbb{Z}/N\mathbb{Z})^2 \cong E[N])_{/K} \middle| e_N(\phi(1,0), \phi(0,1)) = \zeta_N\right] \cong Y(N)(K),$$

where K is any field extension of $\mathbb{Q}[\zeta_N]$ for a specific primitive Nth root of unity ζ_N, and $[\cdot]$ indicates the set of isomorphism classes of the pairs (E, ϕ). We have $(E, \phi)_{/K} \cong (E', \phi')_{/K}$ if there exists an isomorphism $f : E \to E'$ defined over K such that $f \circ \phi = \phi'$.

Since $\mathbb{A} = \mathbb{A}^{(\infty)} \times \mathbb{R}$, we have $GL_2(\mathbb{A})_+/\mathbb{Q}^\times GL_2(\mathbb{R})_+ = \frac{GL_2(\mathbb{A}^{(\infty)})}{Z(\mathbb{Q})}$, writing $Z(\mathbb{Q})$ for the center of $GL_2(\mathbb{Q})$ isomorphic to $\mathbb{Q}^\times$. Consider τ giving the isomorphism

$$\tau : \frac{GL_2(\mathbb{A}^{(\infty)})}{Z(\mathbb{Q})} \cong \mathrm{Aut}(\mathfrak{K}).$$

The fixed field $\mathfrak{K}^S$ of $\tau(S)$ is an algebraic function field; so, it is the function field of a unique smooth projective curve V_S defined over $k_S = \mathfrak{K}^S \cap \overline{\mathbb{Q}}$ (which is the fixed field of $\iota(\det(S))$ for ι in Theorem 2.12). Therefore the isomorphism τ in Theorem 2.43 gives rise to a tower of algebraic curves $\{V_S\}_S$ defined over k_S indexed by open compact subgroups $S \subset \frac{GL_2(\mathbb{A}^{(\infty)})}{Z(\mathbb{Q})}$. The tower $\{V_S\}_S$ is called Shimura's canonical model of the tower of modular curves (or simply just the tower of modular curves), which classifies elliptic curves with additional structures and is the simplest example of the Shimura varieties we study in the rest of the book. We write $Y_S(\mathbb{C})$ for the image of $\Gamma_S\backslash\mathfrak{H}$ in $V_S(\mathbb{C})$ for $\Gamma_S = PGL_2(\mathbb{Q}) \cap S$ in $\frac{GL_2(\mathbb{A}^{(\infty)})}{Z(\mathbb{Q})}$. Then $V_S - Y_S$ is a set of finite points (called cusps of V_S), and Y_S is an open algebraic curve (a quasi-projective curve) defined over k_S. In other words, V_{S/k_S} is the unique smooth compactification of Y_{S/k_S}.

We may regard V_S as defined over $\mathbb{Q}$ forgetting the requirement of the field of definition to be the algebraic closure of $\mathbb{Q}$ in the function field $k_S(V_S)$. By the strong approximation theorem (i.e., the density of $SL_2(\mathbb{Q})$ in $SL_2(\mathbb{A}^{(\infty)})$; e.g., [MFG] 3.1.2), $GL_2(\mathbb{Q})\backslash GL_2(\mathbb{A}^{(\infty)})/Z(\mathbb{Q})S$ is in bijection with $\mathrm{Gal}(k_S/\mathbb{Q})$ by $g \leftrightarrow \iota(\det(g))$. We write ${}^gS$ for the conjugate $g \cdot S \cdot g^{-1}$. Since $k_S \otimes_{\mathbb{Q}} \mathbb{C} = \prod_{\sigma \in \mathrm{Gal}(k_S/\mathbb{Q})} \mathbb{C}$ by $k \otimes x \mapsto (\sigma(k)x)_\sigma$, we have

$$\mathfrak{K}^S \otimes_{\mathbb{Q}} \mathbb{C} = \prod_{g \in GL_2(\mathbb{Q})\backslash \frac{GL_2(\mathbb{A}^{(\infty)})}{Z(\mathbb{Q})}/S} \mathbb{C}\mathfrak{K}^{{}^gS},$$

which implies

$$\begin{aligned} V_S \times_{\mathbb{Q}} \mathbb{C} &= \bigsqcup_{\sigma \in \mathrm{Gal}(k_S/\mathbb{Q})} V_{S/\mathbb{C}}^{\sigma} = \bigsqcup_{g \in GL_2(\mathbb{Q}) \backslash \frac{GL_2(\mathbb{A}^{(\infty)})}{Z(\mathbb{Q})}/S} V_{gS/\mathbb{C}} \\ V_{S/\mathbb{Q}}(\mathbb{C}) &= \bigsqcup_{g \in GL_2(\mathbb{Q}) \backslash \frac{GL_2(\mathbb{A}^{(\infty)})}{Z(\mathbb{Q})}/S} V_{gS/\mathbb{C}}(\mathbb{C}), \end{aligned} \tag{2.22}$$

where each $V_{gS/\mathbb{C}}(\mathbb{C}) = V_{gS/k_S}(\mathbb{C})$ gives one connected component of $V_{S/\mathbb{Q}}(\mathbb{C})$. Since $V_S(\mathbb{C})$ with finitely many points above ∞ (called cusps) removed is exactly the quotient $Y_S = \Gamma_S \backslash \mathfrak{H}$, we find

$$GL_2(\mathbb{Q})_+ \backslash \left(\frac{GL_2(\mathbb{A}^{(\infty)})}{Z(\mathbb{Q})} \times \mathfrak{H} \right) / S = \bigsqcup_{g \in GL_2(\mathbb{Q}) \backslash \frac{GL_2(\mathbb{A}^{(\infty)})}{Z(\mathbb{Q})}/S} Y_{gS}(\mathbb{C}) = Y_{S/\mathbb{Q}}(\mathbb{C})$$

by $(g, z) \mapsto (z \mod \Gamma_{gS})$. Thus we get the following expression:

$$\begin{aligned} &GL_2(\mathbb{Q})_+ \backslash \left(GL_2(\mathbb{A}^{(\infty)}) \times \mathfrak{H} \right) / Z(\mathbb{Q}) \\ &\quad = \varprojlim_S GL_2(\mathbb{Q})_+ \backslash \left(\frac{GL_2(\mathbb{A}^{(\infty)})}{Z(\mathbb{Q})} \times \mathfrak{H} \right) / S = \varprojlim_S Y_{S/\mathbb{Q}}(\mathbb{C}). \end{aligned} \tag{2.23}$$

We may think of the proalgebraic curve $Y = \varprojlim_S Y_{S/\mathbb{Q}}$ as a model of the pro-Riemann surface $GL_2(\mathbb{Q})_+ \backslash \left(GL_2(\mathbb{A}^{(\infty)}) \times \mathfrak{H} \right) / Z(\mathbb{Q})$. This is the point of view of Deligne ([D1] and [D2]). Then $\varprojlim_S V_S$ is the smooth compactification of the open proalgebraic curve Y. In the rest of this book, we replace the algebraic group $GL(2)_{/\mathbb{Q}}$ by a more general reductive group $G_{/\mathbb{Q}}$ and study the canonical models of $G(\mathbb{Q}) \backslash \left(G(\mathbb{A}^{(\infty)}) \times X \right) / \overline{Z(\mathbb{Q})}$ for the symmetric space X of $G(\mathbb{R})$ and their (smooth and minimal) compactifications. Here $\overline{Z(\mathbb{Q})}$ is the topological closure of the center $Z(\mathbb{Q})$ of $G(\mathbb{Q})$ in $G(\mathbb{A}^{(\infty)})$. In the case of $G = GL(2)_{/\mathbb{Q}}$, we have $\overline{Z(\mathbb{Q})} = Z(\mathbb{Q})$; so, our formulation is consistent. In our study of p-adic automorphic forms on G, the two formulations, one due to Deligne and the other due to Shimura, both play fundamental roles.

Exercises

1. Prove the surjectivity of the mod N map from $SL_2(\mathbb{Z})$ to $SL_2(\mathbb{Z}/N\mathbb{Z})$.
2. Let $\delta = \left(\begin{smallmatrix} m & 0 \\ 0 & n \end{smallmatrix}\right)$. Prove that $\delta^{-1}\Gamma(N)\delta \supset \Gamma(mnN)$.
3. Prove (A1) and (A2).

3

Modular Curves

In this chapter, we translate the result in the latter part of the previous chapter into the language of schemes, and at the same time, we sketch a proof of the vertical control theorem in the p-ordinary case of elliptic modular forms. There are several different proofs of the vertical control theorem:

(1) Through the moduli theory of elliptic curves; this is what we show (Ann. Sci. Ec. Norm. Sup. **19** (1986) and [GME] Chapter 3);
(2) Through the study of topological cohomology groups and Jacobians of modular curves. This way has the advantage of simultaneously producing at the same time Galois representations into $GL_2(\mathbb{I})$, where $\mathbb{I}$ is a quotient of the universal p-ordinary Hecke algebra. The ring $\mathbb{I}$ could be large and may be free of finite type over $W[[X]]$ (Inventiones **85** (1986));
(3) Through the theory of p-adic Eisenstein measures and p-adic Rankin convolution theory. This method was found by A. Wiles in [Wi] and is presented in the elliptic modular case in my book [LFE] in Chapter 7;
(4) As an application of the identification of Hecke algebras and universal Galois deformation rings at many different weights (done by Wiles and Taylor). This method is presented in my book [MFG] 5.3.5.

We follow the first method. Chapter 3 of the book [GME] contains a more down-to-earth description of the proof. This first method seems most effective when we try to generalize the VCT to more general reductive groups (as long as the group admits Shimura varieties). We generalize this proof to automorphic forms on more general Shimura varieties (see Sections 4.1, 4.2, and 8.3).

3.1 Basics of Elliptic Curves over a Scheme

We give a brief description of the theory of the moduli of elliptic curves. Chapter 2 of the book [GME] contains a thorough exposition of the theory.

Studying curves C over a field k in Section 2.4, we have analyzed its set $C(K)$ of K-points for a field extension K/k. The main idea behind the theory in Section 2.4 is to associate with each point $P \in C(K)$ a discrete valuation of the function field $K(C)$. However, to study a curve C over a base ring $\mathcal{B}$, obviously we need to study $C(R)$ directly for an arbitrary $\mathcal{B}$-algebra R. The main point of scheme theory is to consider C as a covariant functor $R \mapsto C(R)$ from the category $\mathcal{B}$-ALG (also written as $ALG_{/\mathcal{B}}$) of $\mathcal{B}$-algebras into the category $SETS$ of sets. We have many interesting functors $F : \mathcal{B}$-$ALG \to SETS$. For example, we may associate with R the set $[E_{/R}]$ of all isomorphism classes of elliptic curves defined over R (or $Spec(R)$). This is a covariant functor since a $\mathcal{B}$-algebra homomorphism $R \xrightarrow{\rho} R'$ induces the base change $E_{/R} \mapsto (E \otimes_R R')$ which is an elliptic curve defined over R'. We can extend this functor to a contravariant functor from the category $\mathcal{B}$-SCH (also written as $SCH_{/\mathcal{B}}$) of schemes over $\mathcal{B}$ to $SETS$ in an obvious manner. The functor is contravariant, because $R \mapsto Spec(R)$ reverses the direction of morphisms. If the functor $R \mapsto [E_{/R}]$ is isomorphic to a functor $R \mapsto \mathfrak{M}(R)$ for a $\mathcal{B}$-scheme $\mathfrak{M}$, we could use algebraic geometry to study the functor (and at the same time we may use the functoriality to study the geometry of the moduli scheme $\mathfrak{M}$). Finding $\mathfrak{M}$ is a classification problem of elliptic curves. Of course, we may add some extra structures ϕ to E to classify pairs (E, ϕ), and if successful, the moduli $\mathfrak{M}'$ of the functor $R \mapsto [(E, \phi)_{/R}]$ constitutes a scheme over $\mathfrak{M}$ by the morphism $(E, \phi) \mapsto E$. We study this type of problem in this and the following sections. In this section, we summarize necessary facts on elliptic curves defined over a scheme.

3.1.1 Definition of Elliptic Curves

For a given scheme S, a proper smooth curve $f : E \to S$ is called an *elliptic curve* if it satisfies the following three conditions:

(E1) E has a section $\mathbf{0} = \mathbf{0}_E \in E(S)$ (thus $f \circ \mathbf{0} = 1_S$);

(E2) $\dim_S E = 1$, and E is geometrically connected (this means that each geometric fiber of E over an algebraically closed field is connected and of dimension 1);

(E3) $f_*\Omega_{E/S}$ (equivalently $R^1 f_* \mathcal{O}_E$ by Grothendieck–Serre duality) is locally free of rank 1 (genus = 1).

Compare this definition to the one already given in Section 2.5. There is no harm in assuming that S is connected, as we do from now on. For any S-scheme $\phi : T \to S$, the fibered product $E_T = E \times_S T$ is again an elliptic curve with the zero section $\mathbf{0}_T = \mathbf{0}_E \times 1_T$. For two elliptic curves E and E' over S, an S-morphism $h : E \to E'$ is always supposed to take $\mathbf{0}_E$ to $\mathbf{0}_{E'}$.

3.1.2 Cartier Divisors

A closed subscheme $D \subset E$ is called an effective Cartier divisor (relative to S) on E if $f_*\mathcal{O}_D = f_*(\mathcal{O}_E/I(D))$ given by an invertible sheaf of ideals

$I(D) \subset \mathcal{O}_E$ is S-flat (thus locally free). We define $\mathcal{L}(D) = I(D)^{-1}$ and put $\deg(D) = \deg(\mathcal{L}(D)) = \operatorname{rank}_S f_*(\mathcal{O}_D)$. In particular, the $\mathbf{0}$ section gives rise to a divisor $[\mathbf{0}]$ of degree 1 given by $\mathcal{O}_{[\mathbf{0}]} \overset{\mathbf{0}^\#}{\cong} \mathcal{O}_S$. We then consider $I(m[\mathbf{0}]) = I([\mathbf{0}])^m$ and $\mathcal{L}(m[\mathbf{0}]) = I(m[\mathbf{0}])^{-1}$ for $m \in \mathbb{Z}$. The line bundle $\mathcal{L}(m[\mathbf{0}])$ can be regarded as the sheaf of meromorphic functions on E smooth outside $\mathbf{0}$ and with possible singularity at $\mathbf{0}$ having a pole of order $\leq m$.

If $S = Spec(k)$ for an algebraically closed field k, k-rational effective divisors can be identified with positive linear combinations of points on $E(k)$. We have $\deg(\sum_P m_P[P]) = \sum_P m_P$. We can thus think of the group $\operatorname{Div}(E/k)$ of all formal linear combinations (including negative coefficients) of points on E. Then $\deg : \operatorname{Div}(E/k) \to \mathbb{Z}$ is a well-defined homomorphism given by the above formula. In particular, for any divisor $D \in \operatorname{Div}(E/k)$, we have $\mathcal{L}(D) = \mathcal{L}(D_+) \otimes \mathcal{L}(D_-)^{-1}$ writing $D = D_+ - D_-$ for effective divisors D_+ and D_-, and we can verify $\deg(\mathcal{L}(D)) = \deg(D)$. By Abel's theorem (Theorem 2.36), any invertible sheaf $\mathcal{L}$ on $E_{/k}$ is isomorphic to $\mathcal{L}(D)$ for a divisor D, and we define $\deg(\mathcal{L}) = \deg(\mathcal{L}(D))$ if E is defined over an algebraically closed field k.

3.1.3 Picard Schemes

For any scheme X, we define $\operatorname{Pic}(X)$ as the set of all isomorphism classes of invertible sheaves on X. The association $X \mapsto \operatorname{Pic}(X)$ is a contravariant functor by the pullback of invertible sheaves, and $\operatorname{Pic}(X)$ is actually a group by tensor product.

Let $E \overset{f}{\to} S$ be an elliptic curve. We define, for each S-scheme $\phi : T \to S$,

$$\operatorname{Pic}_{E/S}(T) = \operatorname{Pic}(E_T)/f_T^* \operatorname{Pic}(T).$$

We can extend the degree map to $\deg : \operatorname{Pic}_{E/S}(T) \to \mathbb{Z}^{\pi_0(T)}$ for the set of connected components $\pi_0(T)$. Indeed, for any algebraically closed field k and a geometric point $s : Spec(k) \hookrightarrow T$, the fiber $E(s) = E \times_{S,\phi\circ s} s = E_T \times_T s$ is an elliptic curve over the field k and $\deg(\mathcal{L}) = \deg(\mathcal{L}(s))$ for the pull back $\mathcal{L}(s)$ at s, which is well-defined independently of the choice of s in a connected component. We embed $\mathbb{Z}$ in $\mathbb{Z}^{\pi_0(T)}$ diagonally, and we define

$$\operatorname{Pic}^r_{E/S}(T) = \left\{ \mathcal{L} \in \operatorname{Pic}_{E/S}(T) \middle| \deg(\mathcal{L}) = r \right\}.$$

Here is a version of Abel's theorem (e.g., Section 2.5.5 and [GME] 2.2.2):

(Abel) $\operatorname{Pic}^r_{E/S}(T) \cong E(T) = \operatorname{Hom}_S(T, E)$ by $\mathcal{L}([P]) \otimes \mathcal{L}([\mathbf{0}])^{r-1} \leftrightarrow P$.

Thus an elliptic curve is a group scheme with the identity $\mathbf{0}$ (see [GME] 1.6 for the basics of group schemes). If $\phi : C \to C'$ is a nonconstant S-morphism of two smooth geometrically connected curves, $|\phi^{-1}(s')|$ is constant for geometric points s' of C'; that is, $\phi_* \mathcal{O}_C$ is locally free of finite rank (e.g., [GME] Lemma 2.6.1). We write this number as $\deg(\phi)$. If the base scheme S

is a spectrum of a field, the valuation ring $\mathcal{V}_{s'}$ associated with a closed point $s' \in C'$ is given by the stalk $\mathcal{O}_{C',s'}$; so, this definition of $\deg(\phi)$ is consistent with the one given for algebraic curves over a field just after Corollary 2.26. Thus $\bigwedge^{\deg(\phi)} \phi_* \mathcal{O}_C$ is an invertible sheaf on C'. If $\phi : E \to E'$ is an S-morphism of elliptic curves, by our convention, ϕ takes $\mathbf{0}_E$ to $\mathbf{0}_{E'}$ and, hence, at the side of the Picard scheme, it is just $\mathcal{L} \mapsto \bigwedge^{\deg(\phi)} \phi_* \mathcal{L}$; so, obviously ϕ is a homomorphism of group functors. In particular, a morphism of elliptic curves (taking zero to zero) is a homomorphism of group schemes.

3.1.4 Invariant Differentials

By (E3), for a dense affine open subset $Spec(\mathcal{B})$ of S, $H^0(E, \Omega_{E/\mathcal{B}}) = \mathcal{B}\omega$ for a 1-differential ω. For each point $P \in E(S)$, $T_P : x \mapsto x + P$ gives an automorphism on E. Since we can therefore bring any given cotangent vector at $\mathbf{0}$ isomorphically to a cotangent vector at P, each cotangent vector at $\mathbf{0}$ extends to a global section of $\Omega_{E/S}$. Thus $T_P^* \omega = \omega$ (see [GME] 2.2.3 for a more rigorous proof of this invariance).

3.1.5 Classification Functors

An important fact from functorial algebraic geometry is that we can associate, with each S-scheme X, a contravariant functor $\underline{X} : S\text{-}SCH \to SETS$ such that $\underline{X}(T) = \mathrm{Hom}_S(T, X)$. This association is fully faithful; in other words, writing CTF for the category of contravariant functors from S-schemes to $SETS$, we have

$$\mathrm{Hom}_S(X, Y) \cong \mathrm{Hom}_{CTF}(\underline{X}, \underline{Y})$$

via $(X \xrightarrow{\phi} Y) \mapsto (\phi(T) : \underline{X}(T) \to \underline{Y}(T))$ given by $\phi(T)(T \xrightarrow{f} X) = \phi \circ f$. This is intuitively clear because an algebraic variety is just a function associating to each ring R its R-integral points $X(R) = \underline{X}(Spec(R))$. The verification of this is left to the reader as an exercise (the inverse is given by $\mathrm{Hom}_{CTF}(\underline{X}, \underline{Y}) \ni F \mapsto F(X)(1_X)$ where $F(X) : \underline{X}(X) \to \underline{Y}(X) = \mathrm{Hom}_S(X, Y)$; see, for example, [MFG] 4.1.3 or [GME] 1.4.3).

This fact can be used as follows. Let N be a positive integer. Since $E(T)$ is a group, $x \mapsto N \cdot x$ gives a functorial map $N(T) : E(T) \to E(T)$; so, an endomorphism of elliptic curves $N : E \to E$. We define its kernel $E[N] = E \times_{E,N,\mathbf{0}} S$:

$$\begin{array}{ccc} E[N] & \longrightarrow & E \\ \downarrow & & \downarrow N \\ S & \xrightarrow[\mathbf{0}]{} & E\,. \end{array}$$

It is clear that $E[N](T) = \mathrm{Ker}(N(T))$. It is known that $\deg N = N^2$ (e.g., [GME] Theorem 2.6.2) and if N is invertible over S, $E[N](k) \cong (\mathbb{Z}/N\mathbb{Z})^2$ for all algebraically closed fields k with $Spec(k) \to S$.

We consider the following functor

$$\mathcal{P}'_{\Gamma_1(N)}(R) = \left[(E, P, \omega)_{/R}\right]$$

from the category ALG of $\mathbb{Z}$-algebras into $SETS$, where ω is a nowhere-vanishing invariant differential, P is a point of order exactly N, that is, $m \mapsto mP$ induces an isomorphism $(\mathbb{Z}/N\mathbb{Z})_{/R} \hookrightarrow E$ of group schemes defined over R, and $[\ \ ] = \{\ \ \}/\cong$ is the set of all isomorphism classes of the objects inside the brackets. Here $\mathbb{Z}/N\mathbb{Z}$ as a group functor associates with T the group $(\mathbb{Z}/N\mathbb{Z})^{\pi_0(T)}$. Therefore $\mathcal{O}_{(\mathbb{Z}/N\mathbb{Z})_{/\mathbb{Z}}} = \bigoplus_{\mathbb{Z}/N\mathbb{Z}} \mathbb{Z}$ (and $\mathcal{O}_{(\mathbb{Z}/N\mathbb{Z})_{/R}} = \bigoplus_{\mathbb{Z}/N\mathbb{Z}} R$); so, the structure sheaf of $\mathbb{Z}/N\mathbb{Z}$ is free of finite rank N. Such a group scheme is called a *locally free* group scheme (of rank N).

There is another example of locally free group schemes: starting with the multiplicative group $\mathbb{G}_m$ (as a group functor $\mathbb{G}_m(R) = R^\times$ and as an affine scheme $\mathbb{G}_m = Spec(\mathbb{Z}[t, t^{-1}])$), we consider the kernel μ_N of N (as a group functor $\mu_N(R) = \{\zeta \in R | \zeta^N = 1\}$ and as a scheme $\mu_N = Spec(\mathbb{Z}[t]/(t^N - 1)) = Spec(\mathbb{Z}[(\mathbb{Z}/N\mathbb{Z})])$). Then μ_N is a locally free group scheme of rank N. If $N > 1$, it is nonisomorphic to $\mathbb{Z}/N\mathbb{Z}$, since for any prime p, $\mu_p(\mathbb{F}_p) = \{1\}$ but $(\mathbb{Z}/p\mathbb{Z})(\mathbb{F}_p) = \mathbb{Z}/p\mathbb{Z}$. We consider a version of the functor $\mathcal{P}'_{\Gamma_1(N)}$ defined as follows,

$$\mathcal{P}_{\Gamma_1(N)}(R) = \left[(E, \phi_N : \mu_N \hookrightarrow E[N], \omega)_{/R}\right].$$

3.1.6 Cartier Duality

The two functors $\mathcal{P}_{\Gamma_1(N)}$ and $\mathcal{P}'_{\Gamma_1(N)}$ are isomorphic by Cartier duality (see [GME] 1.7): If G is a locally free group scheme of rank N over S, there exists a group scheme $\widehat{G}_{/S}$ such that $\widehat{G}(T) = \mathrm{Hom}_T(G_T, \mathbb{G}_{m/T}) = \mathrm{Hom}_T(G_T, \mu_{N/T})$, where $\mathbb{G}_{m/S} = \mathbb{G}_m \times S$ and $\mu_{N/S} = \mu_N \times S$ over $Spec(\mathbb{Z})$. We have $\widehat{\widehat{G}} \cong G$ in an obvious manner, and $\widehat{\mathbb{Z}/N\mathbb{Z}} = \mu_N$ by $\zeta(m) = \zeta^m$ for $\zeta \in \mu_N(R)$ and $m \in (\mathbb{Z}/N\mathbb{Z})(R)$.

Let $E \xrightarrow{f} S$ be an elliptic curve. The section $\mathbf{0} : S \to E$ induces a section of $f^* : \mathrm{Pic}(S) \to \mathrm{Pic}(E)$; so, we have a splitting:

$$\mathrm{Pic}(E_T) = f_T^* \mathrm{Pic}(T) \oplus \mathrm{Ker}(\mathbf{0}_T^*) \quad \text{and} \quad \mathrm{Ker}(\mathbf{0}_T^*) = \mathrm{Pic}_{E/S}(T),$$

regarding $\mathrm{Pic}_{E/S}(T)$ as a set of isomorphism classes of invertible sheaves whose restriction to $\mathbf{0}$ is trivial; that is, $\mathbf{0}^*\mathcal{L}$ is isomorphic to $\mathcal{O}_S$.

Let $P \in \mathrm{Ker}(\pi)$ for a nonconstant S-morphism $\pi : E \to E'$. Then

$$\pi^* : E' \cong \mathrm{Pic}_{E'/S} \to \mathrm{Pic}_{E/S} \cong E$$

is an S-homomorphism. We show that $\mathrm{Ker}(\pi^*) = \widehat{\mathrm{Ker}(\pi)}$. Take $\mathcal{L} \in \mathrm{Ker}(\pi^*)$, and take an open covering $E' = \bigcup_i U_i$ such that $\mathcal{L}|_{U_i} = f_i^{-1}\mathcal{O}_{U_i}$. Since $\mathbf{0}^*\mathcal{L} = \mathcal{O}_S$, we may normalize f_i so that $f_i \circ \mathbf{0}_{E'} = f_j \circ \mathbf{0}_{E'}$ for all $i \neq j$ on $U_i \cap U_j$.

Let $h_i = f_i \circ \pi$; we have $\pi^*\mathcal{L}|_{V_i} = h_i^{-1}\mathcal{O}_{V_i}$ for $V_i = \pi^{-1}(U_i)$. Let $P \in \mathrm{Ker}(\pi)$; then

$$h_i \circ P = f_i \circ \pi \circ P = f_i \circ \mathbf{0}_{E'} = f_j \circ \mathbf{0}_{E'} = h_j \circ P.$$

This implies that the $h_i \circ P$s glue to give a global section $h \circ P \in \Gamma(T, \mathcal{O}_T^\times) = \mathbb{G}_m(T)$, getting a homomorphism $\mathrm{Ker}(\pi^*) \to \widehat{\mathrm{Ker}(\pi)}$, which can be easily verified to be an isomorphism (because twice this operation yields an identity map of $\mathrm{Ker}(\pi^*)$). Since $N^* = N$ as we can see easily, we get $\widehat{E[N]} = E[N]$. Writing the pairing as $\langle\ ,\ \rangle : E[N] \times_S E[N] \to \mu_{N/S}$, we get $\langle \phi(P), Q\rangle = \langle P, \phi^*(Q)\rangle$; so, $\phi \mapsto \phi^*$ is an involution with $\phi^* \circ \phi = \deg(\phi) \geq 0$ (a positive involution; see [GME] 2.6.3). The pairing $\langle \cdot, \cdot \rangle$ is the scheme-theoretic version of the pairing $e_N(\cdot, \cdot)$ we studied in (P1–3) in the proof of Theorem 2.38.

For a given additive level N-structure $\phi_N : \mathbb{Z}/N\mathbb{Z} \hookrightarrow E[N]$, by duality, we get $\pi_N : E[N] \twoheadrightarrow \mu_N$ which has a section ϕ'_N well-determined modulo $C = \phi_N(\mathbb{Z}/N\mathbb{Z})$. Thus $(E/C, \phi'_N : \mu_N \hookrightarrow (E/C)[N], \omega')$ is well-defined as an element of $\mathcal{P}_{\Gamma_1(N)}(R)$, where ω' coincides with ω at the identity (because the projection $E \to E/C$ is a local isomorphism; that is, an étale morphism; see [GME] Proposition 1.8.4). The inverse map: $\mathcal{P} \to \mathcal{P}'$ is given similarly by

$$(E', \phi'_N : \mu_N \hookrightarrow E'[N], \omega') \mapsto (E'' = E'/\mathrm{Im}(\phi'_N), \phi_N : \mathbb{Z}/N\mathbb{Z} \hookrightarrow E''[N], \omega'').$$

Since $(E/C)/\phi'_N(\mu_N) = E/E[N] \cong E$, we have $\mathcal{P}' \cong \mathcal{P} \cong \mathcal{P}'$ and hence equivalence.

3.2 Moduli of Elliptic Curves and the Igusa Tower

As already described, we regard a scheme M over a ring $\mathcal{B}$ as a contravariant functor $\underline{M} : SCH_{/\mathcal{B}} \to SETS$ by $\underline{M}(S) = \mathrm{Hom}_{\mathcal{B}\text{-}SCH}(S, M)$. The scheme $M_{/\mathcal{B}}$ is a coarse modulus of a contravariant functor $\mathcal{F} : SCH_{/\mathcal{B}} \to SETS$ if we have a morphism $f : \mathcal{F} \to \underline{M}$ of functors such that $f(k) : \mathcal{F}(Spec(k)) \cong \underline{M}(Spec(k))$ for any algebraically closed field k and for any other morphism $g : \mathcal{F} \to \underline{M'}$ for a $\mathcal{B}$-scheme M', we have a morphism $M \xrightarrow{\pi} M'$ with $g = \pi \circ f$ (see [GME] 2.3.2 (CM1–3) for more details on coarse moduli schemes). If the above morphism f is an isomorphism of functors, we say that $\mathcal{F}$ is *representable* by $M_{/\mathcal{B}}$, and $M_{/\mathcal{B}}$ is a *fine modulus* scheme (see Section 4.4.1). A fine or coarse modulus of a functor is uniquely determined up to isomorphisms of $\mathcal{B}$-schemes (Yoneda's lemma; see [GME] Lemma 1.4.1).

We now study the scheme $Y_1(N)_{/\mathbb{Z}[\frac{1}{N}]}$ representing $\mathcal{P}_{\Gamma_1(N)}$. This eventually leads us to the vertical control theorems in the elliptic modular cases.

3.2.1 Moduli of Level 1 over $\mathbb{Z}[\frac{1}{6}]$

Hereafter, we assume until Section 3.3 (for simplicity) that 6 is invertible in any algebra we consider (see [GME] 2.6–9 for the theory valid over $\mathbb{Z}$). Let

$(E,\omega)_{/R}$ be a pair of an elliptic curve and a nowhere-vanishing differential. We choose a parameter T at $\mathbf{0}$ so that

$$\omega = (1 + \text{higher terms of } T)dT.$$

By the integral version of the Riemann–Roch theorem (see [GME] 2.1.4), we verify (as in Proposition 2.32) that $\operatorname{rank} H^0(E, \mathcal{L}(m[\mathbf{0}])) = m$ if $m > 0$. Therefore we have two morphisms $x, y : E \to \mathbf{P}^1$ such that

1. x has a pole of order 2 at $\mathbf{0}$ with the leading term T^{-2} in its Laurent expansion in T (removing the constant term by translation);
2. y has a pole of order 3 with leading term $-T^{-3}$.

We repeat here the computation of the Weierstrass equation of E done in Section 2.5.2 more carefully over the ring R as follows. Out of these functions x and y, we can create bases of $H^0(E, \mathcal{L}(m[\mathbf{0}]))$:

- $H^0(E, \mathcal{L}(2[\mathbf{0}])) = R + Rx$, $H^0(E, \mathcal{L}(3[\mathbf{0}])) = R + Rx + Ry$. This implies that x has a pole of order 2 at $\mathbf{0}$ and y has order 3 at $\mathbf{0}$. They are regular outside $\mathbf{0}$;
- From these functions $1, x, y$, we create functions with a pole of order n at $\mathbf{0}$ as follows:

$$\begin{aligned} &n \le 4 : 1, x, y, x^2 \ (\dim = 4)\\ &n \le 5 : 1, x, y, x^2, xy \ (\dim = 5)\\ &n \le 6 : 1, x, y, x^2, xy, x^3, y^2 \ (\dim = 6). \end{aligned}$$

Comparing the leading term of T^{-6}, one sees that the seven sections

$$1, x, y, x^2, xy, x^3, y^2$$

in the space $H^0(E, \mathcal{L}(6[\mathbf{0}]))$ have to be linearly dependent and satisfy the following relation,

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

We can kill in a unique way the terms involving xy and y by a variable change $y \mapsto y + ax + b$. Indeed, by the variable change $y \mapsto y - \frac{a_1}{2}x - \frac{a_3}{2}$, we get a simplified equation:

$$y^2 = x^3 + b_2x^2 + b_4x + b_6.$$

Again a variable change $x \mapsto x - \frac{b_2}{3}$ simplifies the equation to

$$y^2 = x^3 + c_2x + c_3.$$

Since $\mathcal{L}(3[\mathbf{0}])$ is very ample ($\deg(\mathcal{L}(3[\mathbf{0}])) = 3 \ge 2g + 1$), by finally making a variable change $2y \mapsto y$ (so now the T-expansion of y begins with $-2T^{-3}$), we get a unique equation from $(E,\omega)_{/R}$:

$$y^2 = 4x^3 - g_2(E,\omega)x - g_3(E,\omega) \quad \text{for } g_2(E,\omega), g_3(E,\omega) \in R.$$

In other words, $E \subset \mathbf{P}^2_{/R}$ is given by

$$Proj(R[X,Y,Z]/(ZY^2 - 4X^3 + g_2(E,\omega)XZ^2 + g_3(E,\omega)Z^3)).$$

It is easy to see by computation that this equation gives a smooth curve of genus 1 having the origin $\mathbf{0} = \infty = (0,1,0)$ in $\mathbf{P}^2$ if $\Delta = \Delta(E,\omega) = g_2^3 - 27g_3^2 \in R^\times$. We recover the differential ω by $\frac{dx}{y}$. This shows that, writing $\mathcal{R} = \mathbb{Z}[\frac{1}{6}, g_2, g_3, \frac{1}{\Delta}]$ for variables g_2 and g_3,

$$\mathcal{P}_{\Gamma_1(1)}(R) \cong \mathrm{Hom}_{\mathbb{Z}[\frac{1}{6}]-alg}(\mathcal{R}, R) = \mathcal{M}_1(R),$$

where $\mathcal{M}_1 = Spec(\mathcal{R})$ for $\mathcal{R} = \mathbb{Z}[\frac{1}{6}, g_2, g_3, \frac{1}{\Delta}]$. We have the universal elliptic curve and the universal differential $\boldsymbol{\omega}$ given by

$$(\mathbf{E}, \boldsymbol{\omega})_{/\mathcal{M}_1} = \left(Proj(\mathcal{R}[X,Y,Z]/(ZY^2 - 4X^3 + g_2XZ^2 + g_3Z^3)), \frac{dX}{Y} \right).$$

For each pair $(E,\omega)_{/R}$, we have a unique $\varphi \in \mathcal{M}_1(R) = \mathrm{Hom}_S(Spec(R), \mathcal{M}_1)$ $(S = Spec(\mathbb{Z}[\frac{1}{6}]))$ such that

$$(E,\omega)_{/R} \cong \varphi^*(\mathbf{E}, \boldsymbol{\omega}) = (\mathbf{E}, \boldsymbol{\omega}) \times_{\mathcal{M}_1} Spec(R).$$

If we change ω by $\lambda\omega$ for $\lambda \in R^\times = \mathbb{G}_m(R)$, the parameter T will be changed to λT and hence (x,y) is changed to $(\lambda^{-2}x, \lambda^{-3}y)$. Thus $(E,\lambda\omega)_{/R}$ will be defined by

$$(\lambda^{-3}y)^2 = 4(\lambda^{-2}x)^3 - g_2(E,\lambda\omega)(\lambda^{-2}x) - g_3(E,\lambda\omega).$$

This has to be equivalent to the original equation by the uniqueness of the Weierstrass equation, and we have

$$g_\jmath(E,\lambda\omega) = \lambda^{-2\jmath} g_\jmath(E,\omega).$$

Again by the uniqueness of the Weierstrass equation, we find that

$$\mathrm{Aut}((E,\omega)_{/R}) = \{1_E\}$$

as long as 6 is invertible in R.

3.2.2 Moduli of $\mathcal{P}_{\Gamma_1(N)}$

Consider $(E,P,\omega)_{/R}$ for a point $P \in E[\ell](R)$ of order ℓ for a prime ℓ. We have a unique $\varphi \in \mathcal{M}_1(R)$ such that

$$\varphi_E : (E,\omega)_{/R} \cong \varphi^*(\mathbf{E}, \boldsymbol{\omega}) = (\mathbf{E}, \boldsymbol{\omega}) \times_{\mathcal{M}_1} Spec(R).$$

We thus have a commutative diagram:

$$\begin{array}{ccc} \mathbf{E} & \to & \mathcal{M}_1 \\ {\scriptstyle\varphi_E}\uparrow & & \uparrow{\scriptstyle\varphi} \\ Spec(R) \xrightarrow{P} & E & \to Spec(R)\,. \end{array}$$

Then P induces a unique morphism $\varphi_P = \varphi_E \circ P : Spec(R) \to (\mathbf{E}[\ell] - \{0\})$. This shows that, over $\mathbb{Z}[\frac{1}{6\ell}]$,

$$\mathcal{P}_{\Gamma_1(\ell)}(R) \cong \mathcal{P}'_{\Gamma_1(\ell)}(R) \cong (\mathbf{E}[\ell] - \{0\})\,(R).$$

Similarly, over $\mathbb{Z}[\frac{1}{6N}]$,

$$\mathcal{P}_{\Gamma_1(N)}(R) \cong \mathcal{P}'_{\Gamma_1(N)}(R) \cong \left(\mathbf{E}[N] - \bigcup_{N>d|N} \mathbf{E}[d]\right)(R).$$

We put $\mathcal{M}_{\Gamma_1(N)} = \mathbf{E}[N] - \bigcup_{N>d|N} \mathbf{E}[d]$. Thus we have proven

Theorem 3.1 *There is an affine scheme $\mathcal{M}_{\Gamma_1(N)} = Spec(\mathcal{R}_{\Gamma_1(N)})$ defined over $\mathbb{Z}[\frac{1}{6N}]$ such that*

$$\mathcal{P}_{\Gamma_1(N)}(R) \cong \mathcal{P}'_{\Gamma_1(N)}(R) \cong \mathrm{Hom}_{\mathbb{Z}[\frac{1}{6N}]-alg}(\mathcal{R}_{\Gamma_1(N)}, R) = \mathcal{M}_{\Gamma_1(N)}(R)$$

for all $\mathbb{Z}[\frac{1}{6N}]$-algebras R. The scheme $\mathcal{M}_{\Gamma_1(N)}/\mathcal{M}_1$ is an étale covering of degree $\varphi(N)$ for the Euler function φ.

The fact that the covering is étale finite follows from the same fact for $\mathbf{E}[N]$ since $\mathbf{E}[N](k) \cong (\mathbb{Z}/N\mathbb{Z})^2$ for all algebraically closed fields with characteristic not dividing N. Since $\mathcal{M}_1$ is affine, any finite covering of $\mathcal{M}_1$ is affine.

If we find an S-scheme $\mathcal{M}$ for a given contravariant functor $\mathcal{P} : S\text{-}SCH \to SETS$ such that we have an isomorphism of functors $\underline{\mathcal{M}} \cong \mathcal{P}$ for $\underline{\mathcal{M}}(S') = \mathrm{Hom}_S(S', \mathcal{M})$, we say that $\mathcal{P}$ is *representable* (or represented) by an S-scheme $\mathcal{M}$. The scheme $\mathcal{M}$ is called the *moduli scheme* of the functor $\mathcal{P}$. Then the statement of the theorem is equivalent to the representability of the functors $\mathcal{P}_{\Gamma_1(N)}$ and $\mathcal{P}'_{\Gamma_1(N)}$ by an affine $\mathbb{Z}[\frac{1}{6N}]$-scheme $\mathcal{M}_{\Gamma_1(N)}$.

Even if $\mathcal{P}$ is not representable by a scheme, we often find a rough approximation of the moduli called a *coarse moduli* scheme so that $\mathcal{P}(Spec(k)) \cong \mathcal{M}(k)$ for the spectrum of any algebraically closed field k that is an S-scheme. Here we do not give the scheme-theoretic characterization of coarse moduli schemes, only referring to [GME] 2.3.2, though we show many coarse moduli schemes in this book.

3.2.3 Action of $\mathbb{G}_m$

The group scheme $\mathbb{G}_m$ acts on the functor $\mathcal{P}_{\Gamma_1(N)}$ in the following way,

$$(E, \phi_N, \omega)_{/R} \mapsto (E, \phi_N, \lambda\omega)_{/R},$$

for $\lambda \in \mathbb{G}_m(R)$. This induces (by the functoriality described in Section 3.1.5) an action of $\mathbb{G}_m$ on $\mathcal{M}_{\Gamma_1(N)}$ and hence on $\mathcal{R}_{\Gamma_1(N)}$.

Here is a general fact on the action of $\mathbb{G}_m$ (see [GME] 1.6.5). Let X be a $\mathcal{B}$-module for a base ring $\mathcal{B}$. Regard X as a functor from $\mathcal{B}$-ALG to the category of $\mathcal{B}$-modules $\mathcal{B}$-MOD by $\underline{X}(R) = X \otimes_{\mathcal{B}} R$. If a group scheme $G_{/\mathcal{B}}$ has an action coming from the following morphism of functors, $\underline{G} \times \underline{X} \to \underline{X}$ (which induces the set-theoretic action $\underline{G}(R) \times \underline{X}(R) \to \underline{X}(R)$ for each $\mathcal{B}$-algebra R), we call X a *schematic representation* of G. It is known that if X has a schematic action of $\mathbb{G}_{m/\mathcal{B}}$, then

$$X = \bigoplus_{\kappa \in \mathbb{Z}} X[\kappa]$$

such that $X[\kappa] = \{x \in \underline{X} | \lambda \cdot x = \lambda^{\kappa} x\}$; that is, $X[\kappa]$ is the eigenspace for the character $\mathbb{G}_m(\mathcal{B}) \to \mathcal{B}^{\times}$ taking $z \in \mathbb{G}_m(\mathcal{B}) = \mathcal{B}^{\times}$ to z^{κ}.

The action of $\mathbb{G}_{m/\mathcal{B}}$ on $\mathcal{M}_{\Gamma_1(N)}$ gives rise to a schematic action on $\mathcal{R}_{\Gamma_1(N)}$ (because it was defined by functorial action; see Section 3.1.5). Thus we can split

$$\mathcal{R}_{\Gamma_1(N)/R} = \bigoplus_{\kappa \in \mathbb{Z}} \mathcal{R}_{\kappa}(\Gamma_1(N); R),$$

where on $f \in \mathcal{R}_{\kappa}(\Gamma_1(N); R)$, $\mathbb{G}_m$ acts by the character $-\kappa$.

Since $f \in \mathcal{R}_{\kappa}(\Gamma_1(N); \mathcal{B})$ is a functorial morphism,

$$\underline{\mathcal{M}}_{\Gamma_1(N)}(R) = \mathcal{P}_{\Gamma_1(N)}(R) \to \mathbf{A}^1(R) = R,$$

we may regard f as a function of $(E, \phi_N, \omega)_{/R}$ with $f((E, \phi_N, \omega)_{/R}) \in R$ satisfying

(G0) $f((E, \phi_N, \lambda\omega)_{/R}) = \lambda^{-\kappa} f((E, \phi_N, \omega)_{/R})$ for $\lambda \in R^{\times} = \mathbb{G}_m(R)$;
(G1) If $(E, \phi_N, \omega)_{/R} \cong (E', \phi'_N, \omega')_{/R}$, then we have

$$f((E, \phi_N, \omega)_{/R}) = f((E', \phi'_N, \omega')_{/R});$$

(G2) If $\rho : R \to R'$ is a morphism of $\mathcal{B}$-algebras, then we have

$$f((E, \phi_N, \omega)_{/R} \times_R R') = \rho(f((E, \phi_N, \omega)_{/R})).$$

If a graded ring $\mathcal{A} = \bigoplus_j \mathcal{A}_j$ has a unit u of degree 1, we see that $\mathcal{A} = \mathcal{A}_0 \otimes_{\mathbb{Z}} \mathbb{Z}[u, u^{-1}]$ and $Spec(\mathcal{A}) = Spec(\mathcal{A}_0) \times \mathbb{G}_m$ by definition; so, $Proj(\mathcal{A}) = Spec(\mathcal{A})/\mathbb{G}_m = Spec(\mathcal{A}_0)$. If $\mathcal{A}$ has a unit of degree $n > 0$, then we still have

$$Proj(\mathcal{A}) = Proj(\mathcal{A}^{(n)}) = Spec(\mathcal{A}_0)$$

for $\mathcal{A}^{(n)} = \bigoplus_j \mathcal{A}_{nj}$ (see [GME] Lemma 1.3.1). Since $\Delta^{-1} \in \mathcal{R} \subset \mathcal{R}_{\Gamma_1(N)}$, the graded ring $\mathcal{R}_{\Gamma_1(N)}$ has a unit of degree 12, and hence, we have

$$\mathbb{G}_m\backslash\mathcal{M}_{\Gamma_1(N)/\mathcal{B}} = Proj(\mathcal{R}_{\Gamma_1(N)/\mathcal{B}}) \cong Spec(\mathcal{R}_0(\Gamma_1(N);\mathcal{B})) =: Y_1(N)_{/\mathcal{B}}.$$

We consider the functor defined over $\mathbb{Z}[\frac{1}{N}]$-ALG given by

$$\mathcal{E}_{\Gamma_1(N)}(R) = \left[(E, \phi_N : \mu_N \hookrightarrow E[N])_{/R}\right].$$

By definition, $\mathcal{E}_{\Gamma_1(N)} = \mathbb{G}_m\backslash\mathcal{P}_{\Gamma_1(N)}$. Since $Proj(\mathcal{R}_{\Gamma_1(N)})$ gives the quotient by $\mathbb{G}_m$ of $Spec(\mathcal{R}_{\Gamma_1(N)})$ (cf. [GME] Theorem 1.8.2), we conclude

Theorem 3.2 (Shimura, Igusa) *We have an open curve*

$$Y_1(N) = Proj(\mathcal{R}_{\Gamma_1(N)}) = \mathbb{G}_m\backslash\mathcal{M}_{\Gamma_1(N)}$$

defined over $\mathbb{Z}[\frac{1}{6N}]$, *which is locally free of finite rank over* $M_1 = Proj(\mathcal{R}) = \mathbf{P}^1(J) - \{\infty\}$. *For all geometric points* $Spec(k)$ *of* $Spec(\mathbb{Z}[\frac{1}{6N}])$, *we have* $\mathcal{E}_{\Gamma_1(N)}(k) = \left[(E, \phi_N)_{/k}\right] = Y_1(N)(k)$. *The above assertion holds for any* $\mathbb{Z}[\frac{1}{6N}]$-*algebras* R *in place of algebraically closed fields* k *if* $N \geq 4$.

Here a "geometric point" means that k is an algebraically closed field. This theorem over $\mathbb{Z}[\frac{1}{M}]$ for an unspecified M with $N|M$ was proved by Shimura in the late 1950s (see [CPS] I [57a]), and the result over $\mathbb{Z}[\frac{1}{N}]$ was proved by Igusa in [I] soon after Shimura's work.

Returning to the classical setting, we define an arithmetic subgroup $\Gamma_1(N)$ of $SL_2(\mathbb{Q})$ by

$$\Gamma_1(N) = \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z}) \middle| c \equiv 0 \mod N,\ a \equiv d \equiv 1 \mod N \right\}.$$

Since $SL_2(\mathbb{Z})$ acts on the upper half complex plane $\mathfrak{H}$ discretely by $z \mapsto \frac{az+b}{cz+d}$, we can make a quotient Riemann surface $\Gamma_1(N)\backslash\mathfrak{H}$ (see [IAT] Chapter 1). By the association: $z \mapsto (E_z(\mathbb{C}) = \mathbb{C}/(\mathbb{Z}z+\mathbb{Z}), P_z = \frac{1}{N} \in E_z(\mathbb{C}))$, it is well known that $\Gamma_1(N)\backslash\mathfrak{H}$ classifies all elliptic curves with a point P of order N over $\mathbb{C}$ (cf. [IAT] Chapter 4 and [GME] 2.4); so, we conclude

$$Y_1(N)(\mathbb{C}) = \Gamma_1(N)\backslash\mathfrak{H}.$$

Thus $Y_1(N)(\mathbb{C})$ is an open Riemann surface.

3.2.4 Compactification

For any $\mathbb{Z}[\frac{1}{6}]$-algebra R, we put

$$G(R) = R[g_2, g_3] = \mathbb{Z}\left[\frac{1}{6}, g_2, g_3\right] \otimes R.$$

Let $G_{\Gamma_1(N)}(\mathbb{Z}[\frac{1}{6N}])$ be the integral closure of the ring $G(\mathbb{Z})$ in the graded ring $\mathcal{R}_{\Gamma_1(N)/\mathbb{Z}[\frac{1}{6N}]}$. To see that $G_{\Gamma_1(N)}(\mathbb{Z}[\frac{1}{6N}])$ is a graded ring, we write $\widetilde{r}$ for the nontrivial homogeneous projection of highest degree of $r \in \mathcal{R}_{\Gamma_1(N)/\mathbb{Z}[\frac{1}{6N}]}$.

If $r \in \mathcal{R}_{\Gamma_1(N)/\mathbb{Z}[\frac{1}{6N}]}$ is integral over $G(\mathbb{Z})$, r satisfies an equation $P(X) = X^n + a_1X^{n-1} + \cdots + a_n = 0$ with $a_j \in G(\mathbb{Z})$. Then $\widetilde{r}$ satisfies $\widetilde{P}(X) = X^n + \widetilde{a}_1X^{n-1} + \cdots + \widetilde{a}_n = 0$, and $\widetilde{r}$ is integral over $G(\mathbb{Z})$. Then by induction of the degree of $\widetilde{r}$, we see that $G_{\Gamma_1(N)}(\mathbb{Z}[\frac{1}{6N}])$ is graded (cf. [BCM] V.1.8).

We put for any $\mathbb{Z}[\frac{1}{6N}]$-algebra R,

$$G_{\Gamma_1(N)}(R) = G_{\Gamma_1(N)}\left(\mathbb{Z}\left[\frac{1}{6N}\right]\right) \otimes R = \bigoplus_{k=0}^{\infty} G_k(\Gamma_1(N); R).$$

We then define $X_1(N)_{/R} = Proj(G_{\Gamma_1(N)}(R))$. By definition, $X_1(N)$ is the normalization of $Proj(G) = Proj(G^{(12)}) = \mathbf{P}^1(J)$ ($J = \frac{(12g_2)^3}{\Delta}$) for $G^{(12)} = \bigoplus_{k=0}^{\infty} G_{12k}$ in $Y_1(N)$. As classically known, J^{-1} has a q-expansion starting with q, that is, $J^{-1} \in q\mathbb{Z}[[q]]$ (see [IAT] (4.6.1)). Thus the completion of the local ring of $\mathbf{P}^1(J)$ at the cusp ∞ is isomorphic to $\mathbb{Z}[\frac{1}{6}][[q]]$. Moreover, we have the Tate curve (e.g., [GME] 2.5):

$$\mathrm{Tate}(q) = Proj(\mathbb{Z}[[q]][\frac{1}{6}][X, Y, Z]/(ZY^2 - 4X^3 + g_2(q)XZ^2 + g_3(q)Z^3)),$$

which extends the universal curve over $\mathbf{P}^1(J) - \{\infty\}$ to $\mathbf{P}^1(J)$ locally at the cusp ∞.

Since $\mathrm{Tate}(q)(R[[q]]) \supset (R[[q]]^\times)/q^{\mathbb{Z}}$ (see [GME] Theorem 2.5.1), we may regard $\mathrm{Tate}(q)$ as the algebraization of the formal quotient $\widehat{\mathbb{G}}_{m/\mathbb{Z}[[q]]}/q^{\mathbb{Z}}$ of the formal multiplicative group $\widehat{\mathbb{G}}_m$; so, it has a canonical level structure

$$\phi_N^{can} : \mu_N \hookrightarrow \widehat{\mathbb{G}}_m \twoheadrightarrow \mathrm{Tate}(q).$$

The Tate couple $(\mathrm{Tate}(q), \phi_N^{can})_{/\mathbb{Z}[[q]]}$ is an elliptic curve over $\mathbb{Z}[[q]][q^{-1}]$; so, by the universality of $Y_1(N)$, we have a morphism

$$\iota_\infty : Spec(\mathbb{Z}[\frac{1}{N}][[q]][q^{-1}]) \to Y_1(N).$$

Since we may regard the Tate curve as a universal formal deformation of a stable curve of genus 1 (with the level structure ϕ_N^{can}) centered at the $\mathbb{Z}[\frac{1}{N}]$-point represented by an ideal (q) of $\mathbb{Z}[\frac{1}{N}][[q]]$ ([GME] 2.5.2–3), the morphism ι_∞ is an infinitesimal isomorphism centered at the cusp ∞ (by the universality of the $Y_1(N)$ and the universality of the Tate curve). Since $X_1(N)$ is the normalization of $\mathbf{P}^1(J)$ in $Y_1(N)$, we conclude that the formal completion along the cusp ∞ on $X_1(N)$ is canonically identified with $R[[q]]$ by ι_∞. Replacing the level structure ϕ_N^{can} by $\phi_N^{can} \circ \alpha$ for $\alpha \in SL_2(\mathbb{Z}/N\mathbb{Z})$, basically by the same argument, the local ring at the cusp $\alpha(\infty)$ of $X_1(N)_{/R}$ is given by $R[\mu_d][[q^{1/d}]]$ for a suitable divisor $d|N$. We need to extend the scalar to $R[\mu_d][[q^{1/d}]]$ because the Tate curve $(\mathrm{Tate}(q), \phi_N^{can} \circ \alpha)$ is only defined over $R[\mu_d][[q^{1/d}]]$ for a suitable divisor $d|N$ dependent on the choice of α. This point is slightly more technical, and we refer the reader to a more detailed

account, which can be found in [AME] Chapter 10 and [GME] 3.1.1. Thus $X_1(N)$ is smooth at the cusps, and moreover $f \in G_k(\Gamma_1(N); R)$ is a function of (E, ϕ_N, ω) satisfying (G0–2) and

(G3) $f(\mathrm{Tate}(q), \phi_N, \omega) \in R[\zeta_N][[q^{1/N}]]$ for any choice of ϕ_N and ω.

Since $\Gamma_1(N)\backslash(\mathfrak{H} \cup \mathbf{P}^1(\mathbb{Q}))$ is a smooth compact Riemann surface and is the normalization of $\mathbf{P}^1(J)$ in $Y_1(N)(\mathbb{C})$ ([IAT] Chapter 1), we conclude

$$X_1(N)(\mathbb{C}) = \Gamma_1(N)\backslash(\mathfrak{H} \cup \mathbf{P}^1(\mathbb{Q})).$$

The space $G_k(\Gamma_1(N); \mathbb{C})$ is the classical space of modular forms on $\Gamma_1(N)$ of weight k. Since $\mathrm{Tate}(q)(R[[q]]) \supset (R[[q]]^\times)/q^{\mathbb{Z}}$, we may consider $\mathrm{Tate}(q)$ to be a "quotient" of $\mathbb{G}_{m/\mathbb{Z}[[q]]}/q^{\mathbb{Z}}$; so, it has a canonical level structure $\phi_N^{can} : \mu_N \hookrightarrow \mathbb{G}_m \twoheadrightarrow \mathrm{Tate}(q)$ and a canonical differential ω_{can} induced by $\frac{dt}{t}$ identifying $\mathbb{G}_m = Spec(\mathbb{Z}[t, t^{-1}])$. In particular,

$$f(q) = f(\mathrm{Tate}(q), \phi_N^{can}, \omega_{can}) = \sum_{n=0}^{\infty} a(n; f)q^n$$

coincides with the Fourier expansion of f at infinity if $f \in G_k(\Gamma_1(N); \mathbb{C})$.

3.2.5 Moduli of $\Gamma(N)$-Level Structure

We consider the following classification problem of level $\Gamma(N)$,

$$\mathcal{P}_{\Gamma(N)}(R) = \left[(E, \omega, \phi_N : (\mathbb{Z}/N\mathbb{Z})^2 \cong E[N]) \middle| \langle \phi_N(1,0), \phi_N(0,1)\rangle = \zeta_N\right],$$

for all $\mathbb{Z}[\frac{1}{6N}, \zeta_N]$-algebras R, where $\zeta_N = \exp(2\pi i \frac{1}{N})$. Writing f for the map $\langle \cdot, \cdot \rangle : \mathbf{E}[N] \times_{\mathcal{M}_1} \mathbf{E}[N] \to \mu_N$ as a morphism of group schemes over $\mathbb{Z}[\frac{1}{N}]$, it is obvious that $\mathcal{P}_{\Gamma(N)}$ is represented over $\mathbb{Z}[\frac{1}{6N}, \zeta_N]$ by the closed affine subscheme $\mathcal{M}_{\Gamma(N)/\mathbb{Z}[\frac{1}{6N}, \zeta_N]} = f^{-1}(\zeta_N)$ of $\mathbf{E}[N] \times_{\mathcal{M}_1} \mathbf{E}[N]$. Writing $\mathcal{M}_{\Gamma(N)} = Spec(\mathcal{R}_{\Gamma(N)})$ for a graded algebra $\mathcal{R}_{\Gamma(N)}$, we define an affine curve

$$Y(N)_{/\mathbb{Z}[\frac{1}{6N}, \zeta_N]} = Proj(\mathcal{R}_{\Gamma(N)}) = Spec(\mathcal{R}_0(\Gamma(N))) = \mathcal{M}_{\Gamma(N)}/\mathbb{G}_m,$$

which represents the following functor,

$$\mathcal{E}_{\Gamma(N)}(R) = \left[(E, \phi_N : (\mathbb{Z}/N\mathbb{Z})^2 \cong E[N]) \middle| \langle \phi_N(1,0), \phi_N(0,1)\rangle = \zeta_N\right],$$

if $N \geq 3$ (see [GME] Theorem 2.6.8). Here $\mathcal{R}_0(\Gamma(N))$ is the degree 0-component of the graded algebra $\mathcal{R}_{\Gamma(N)}$. Taking the integral closure $G_{\Gamma(N)}$ of the graded algebra $G(\mathbb{Z}[\frac{1}{6N}, \zeta_N])$ in $\mathcal{R}_{\Gamma(N)}$, we define the compactification $X(N) = Proj(G_{\Gamma(N)})$ over $\mathbb{Z}[\frac{1}{6N}, \zeta_N]$. Again we have well-defined q-expansion at ∞.

We can let a constant group $\alpha \in SL_2(\mathbb{Z}/N\mathbb{Z})$ act on $Y(N)$ (and hence on $X(N)$) by $(E, \phi) \mapsto (E, \phi \circ \alpha)$. Since $\langle \phi \circ \alpha(1,0), \phi \circ \alpha(0,1)\rangle = \zeta_N^{\det(\alpha)}$, the same

action of $\alpha \in GL_2(\mathbb{Z}/N\mathbb{Z})$ induces an automorphism of $Y(N)$ (and $X(N)$) regarded as schemes over $\mathbb{Z}[\frac{1}{6N}]$ (not over $\mathbb{Z}[\frac{1}{6N}, \zeta_N]$), which coincides with the Galois action $\zeta_N \mapsto \zeta_N^{\det(\alpha)}$ on $\mathbb{Z}[\zeta_N]$. It is easy to see that $\mathbb{Q}[\zeta_N](X(N)) = \mathfrak{K}_N$ in Corollary 2.41 and the action of $GL_2(\mathbb{Z}/N\mathbb{Z})$ introduced here coincides with the identification $\mathrm{Gal}(\mathfrak{K}_N/\mathbb{Q}(J)) \cong \mathrm{Gal}(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$ in Theorem 2.40.

3.2.6 Hasse Invariant

Let R be a ring of characteristic p and (E, ω) be an elliptic curve over $S = Spec(R)$. On each affine open subset $U = Spec(\Gamma(U, \mathcal{O}_E))$ in E, the Frobenius endomorphism $x \mapsto x^p$ induces a morphism $F_{abs} : U \to U$. These glue each other to the absolute Frobenius endomorphism $F_{abs} : E_{/R} \to E_{/R}$. Note here that F_{abs} acts nontrivially on the coefficient ring R. We can define the relative Frobenius map $E \to E^{(p)} = E \times_{S, F_{abs}} S$ by $F_{abs} \times_S f$ for the structure morphism $f : E \to S$. This relative Frobenius map is the classical map taking homogeneous coordinates on E to their p-powers.

Let $\mathcal{T}_{E/S}$ be the relative tangent bundle. By definition, its sections are derivations; in particular, the space of global sections $H^0(E, \mathcal{T}_{E/S})$ is the R-dual of $H^0(E, \Omega_{E/S})$, and $H^0(E, \mathcal{T}_{E/S})$ is spanned by the dual base $\eta = \eta(\omega)$ of ω. One can identify $H^0(E, \mathcal{T}_{E/S})$ with the module of $\mathcal{O}_S$-derivations $Der_{\mathcal{O}_S}(\mathcal{O}_{E,\mathbf{0}}, \mathcal{O}_S)$ (cf. [GME] 1.5.1). For each derivation D of $\mathcal{O}_{E,\mathbf{0}}$, by the Leibniz formula, we have

$$D^p(xy) = \sum_{j=0}^{p} \binom{p}{j} D^{p-j}xD^j y = xD^p y + yD^p x.$$

Thus D^p is again a derivation. The association $D \mapsto D^p$ induces an F_{abs}-linear endomorphism F^* of $\mathcal{T}_{E/S}$. Then we define $H(E, \omega) \in R$ by $F^*\eta = H(E, \omega)\eta$. Since $\eta(\lambda\omega) = \lambda^{-1}\eta(\omega)$, we see

$$\begin{aligned} H(E, \lambda\omega)\eta(\lambda\omega) &= F^*\eta(\lambda\omega) = F^*(\lambda^{-1}\eta(\omega)) = \lambda^{-p}F^*\eta(\omega) \\ &= \lambda^{-p}H(E, \omega)\eta(\omega) = \lambda^{-p}H(E, \omega)\lambda\eta(\lambda\omega) = \lambda^{1-p}H(E, \omega)\eta(\lambda\omega). \end{aligned}$$

Thus we get

$$H(E, \lambda\omega) = \lambda^{1-p}H(E, \omega).$$

Then H is a modular form of weight $p-1$ defined over $\mathbb{F}_p$:

$$H(E, \omega) \in G_{p-1}(\Gamma_1(1); \mathbb{F}_p).$$

We compute $H(\mathrm{Tate}(q), \omega_{can})$ for $\omega_{can} = \frac{dt}{t}$, writing the coordinate of $\mathbb{G}_m$ as t (i.e., $\mathbb{G}_m = Spec(\mathbb{Z}[t, t^{-1}])$). The dual of $\frac{dt}{t}$ is given by $D = t\frac{d}{dt}$. The action of F keeps D intact, because $D(t) = t$ (so $D^p(t) = t$). On the tangent space, F acts as identity, and hence $H(\mathrm{Tate}(q), \omega_{can}) = 1$.

An important fact (cf. [GME] Proposition 2.9.1) is:

$$H(E,\omega)=0 \iff E \text{ is supersingular.}$$

This is because:

1. If $E_{/\overline{\mathbb{F}}_p}$ is ordinary, then $E[p] \cong \mu_p \times (\mathbb{Z}/p\mathbb{Z})$ over $\overline{\mathbb{F}}_p$;
2. $\mu_p = Spec(\mathbb{F}_p[t]/(t^p-1))$ shares the tangent space with $\mathbb{G}_{m/\mathbb{F}_p}$, because they are both of dimension 1 infinitesimally;
3. $F^2 = p$ up to units in the supersingular case.

The zero locus of a nonzero section of a line bundle is an effective divisor; hence, for any algebraically closed field k of characteristic p for $p \nmid N$, the points in $X_1(N)(k)$ carrying supersingular elliptic curves are finitely many and algebraic over $\mathbb{F}_p$.

3.2.7 Igusa Curves

Let $W = \mathbb{Z}_p$ and $W_m = W/p^m W$. Fix N with $p \nmid N$. We have a lift of the Hasse invariant in $G_{p-1}(\Gamma_1(1); \mathbb{Z}_p)$, which is the Eisenstein series $\mathcal{E}$ normalized so that $a(0,\mathcal{E}) = 1$. The standard Eisenstein series E_k of weight k on $SL_2(\mathbb{Z})$ has the following q-expansion ([LFE] Chapter 5),

$$E_k(q) = \frac{\zeta(1-k)}{2} + \sum_{n=1}^{\infty}\left(\sum_{0<d|n} d^{k-1}\right) q^n.$$

Define $\mathcal{E}(q) = \frac{2}{\zeta(2-p)} E_{p-1}$. By the Von Staut theorem, the q-expansion $\mathcal{E}(q)$ of $\mathcal{E}$ is congruent to 1 modulo p; so, $\mathcal{E} \mod p$ coincides with H. Let $(\mathbf{E}, \phi_N)_{/M}$ be the genus 1 semi-stable curve (completed by appropriate Tate curves at the cusps) over $M = X_1(N)_{/W}$. Let

$$M_m = X_1(N)_{/W_m} = X_1(N) \times_W W_m.$$

Define $S_m \subset M_m$ by the maximal open subscheme of M_m on which $\mathcal{E}$ is invertible, and symbolically write $S_m = M_m\left[\frac{1}{\mathcal{E}}\right]$. Similarly we write $S_0 = M\left[\frac{1}{\mathcal{E}}\right]$ for the maximal open subscheme of M on which $\mathcal{E}$ is invertible. The congruence $\mathcal{E} \equiv \mathcal{E}' \equiv H \mod p$ tells us that $\mathcal{E}' = \mathcal{E}\cdot(1+n)$ for n nilpotent; so, $\mathcal{E}$ and $\mathcal{E}'$ are invertible at the same time. Thus the scheme S_m does not depend on the choice of the lift $\mathcal{E}$ as long as p is nilpotent in the base ring. We write S_∞ for the formal completion $\varprojlim_m S_m$ of S_0 along S_1.

Since we have defined $X_1(N)$ by $Proj(G_{\Gamma_1(N)})$, the invertible sheaf $\underline{\omega}^k$ ($k > 0$) associated with the kth graded piece is ample. To see for which k, $\underline{\omega}^k$ becomes very ample, we recall that an invertible sheaf of degree $\geq 2g+1$ over a curve of genus g is very ample by the Riemann–Roch theorem (see [GME] Proposition 2.1.4). Computing the genus of $X_1(N)$ (e.g., [GME] Theorem 3.1.2), the invertible sheaf $\underline{\omega}^k_{/R}$ corresponding to $G_k(\Gamma_1(N); R)$ is very ample if $k \geq 2$ and $N \geq 4$ (or $k > 2$). Thus S_m is affine, and $S_m = Spec(V_{m,0})$ for a W_m-flat algebra $V_{m,0}$.

Let W_m-$pALG$ be the category of p-adically complete W_m-algebras R with p-adically continuous morphisms. When we write W_∞, it means $W = \varprojlim_m W_m$. Here the words "p-adically complete" mean that $R = \varprojlim_n R/p^n R$. We consider the functors $\mathcal{E}_\alpha^{ord}, \mathcal{E}'^{ord}_\alpha : W_m\text{-}pALG \to SETS$ given by

$$\mathcal{E}'^{ord}_\alpha(R) = \left[(E, P, \phi_N)_{/R}\right] \quad \text{and} \quad \mathcal{E}_\alpha^{ord}(R) = \left[(E, \mu_{p^\alpha} \hookrightarrow E[p^\alpha], \phi_N)_{/R}\right],$$

where P is a point of order p^α. For a given test triple $(E, P, \phi_N)_{/R}$, we have a unique commutative diagram (up to isomorphisms)

$$\begin{array}{ccc} E & \xrightarrow{\iota} & \mathbf{E} \\ \downarrow & & \downarrow \\ Spec(R) & \longrightarrow & Y_1(N) \end{array}$$

and $\iota(P) \in \left(\mathbf{E}[p^\alpha]^{et} - \mathbf{E}[p^{\alpha-1}]^{et}\right)(R)$. The association $(E, P, \phi_N)_{/R} \mapsto \iota(P)$ gives rise to the following isomorphism,

$$\mathcal{E}^{ord}(R) \cong \mathcal{E}'^{ord}(R) \cong \left(\mathbf{E}[p^\alpha]^{et} - \mathbf{E}[p^{\alpha-1}]^{et}\right)_{/S_m}(R),$$

if m is finite. Here $\mathbf{E}[p^\alpha]^{et}$ is the maximal étale quotient of $\mathbf{E}[p^\alpha]$; thus, writing $\mathbf{E}[p^\alpha]$ as a relative affine spectrum $Spec_{S_m}(R)$ (cf. [GME] 1.5.4) of a sheaf of $\mathcal{O}_{S_m}$-algebras R, we have $\mathbf{E}[p^\alpha]^{et} = Spec_{S_m}(R^{et})$, where R^{et} is the maximal subalgebra of R étale over $\mathcal{O}_{S_m}$ (see [T] (1.4)). We write $T_{m,\alpha/S_m} = \left(\mathbf{E}[p^\alpha]^{et} - \mathbf{E}[p^{\alpha-1}]^{et}\right)_{/S_m}$, which is an étale Galois covering with the Galois group $(\mathbb{Z}/p^\alpha\mathbb{Z})^\times$. Then the scheme $T_{m,\alpha}$ represents the functors $\mathcal{E}'^{ord}_\alpha$ and $\mathcal{E}_\alpha^{ord}$ over W_m. The curve $T_{m,\alpha/W_m}$ is called the Igusa curve of level Np^α. We also define a procurve $T_{m,\infty/W_m} = \varprojlim_\alpha T_{m,\alpha}$ which represents the functor $\mathcal{E}_\infty^{ord}$ over W_m.

By a result of Igusa (e.g. [GME] 2.9.3), the curve $T_{1,\alpha/S_1}$ $(\alpha = 1, 2, \ldots, \infty)$ is irreducible (which we prove in the following subsection).

3.2.8 Irreducibility of Igusa Curves

Since $T_{m,\alpha}$ has the same topological space as $T_{1,\alpha}$, we may assume $m = 1$ in order to prove the irreducibility of $T_{m,\alpha}$; namely, we prove the following theorem.

Theorem 3.3 (Igusa) *Let p be a prime outside N and $\overline{\mathbb{F}}_p$ be an algebraic closure of $\mathbb{F}_p$. Then the Igusa curve $T_{\alpha/\overline{\mathbb{F}}_p} := T_{1,\alpha}$ over $X_1(N)_{/\overline{\mathbb{F}}_p}$ is irreducible for all α.*

This theorem was first shown by Igusa in [I] using the fact that the inertia group at each supersingular point of $X_1(N)_{/\overline{\mathbb{F}}_p}$ in $\mathrm{Gal}(T_\alpha/X_1(N))$ is equal to the full group $(\mathbb{Z}/p^\alpha\mathbb{Z})^\times$. We prove this theorem assuming $p > 3$, since we

have constructed modular curves only over $\mathbb{Z}[\frac{1}{6N}]$. Our proof is new and can be generalized easily to more general Shimura varieties (see Theorem 4.21, Theorem 6.27, and Corollary 8.17).

Proof. We write $T_\alpha = Spec(V_\alpha)$ for an étale extension of V_0 over $\mathbb{F}_p$. For two integers $0 < N|N'$ prime to p, T_α over $X_1(N')$ is a finite covering of T_α over $X_1(N)$, the irreducibility of the Igusa curve over $X_1(N')$ implies that of the Igusa curve over $X_1(N)$. Thus we may assume that N is large so that $X_1(N)$ is smooth over $\mathbb{Z}[\frac{1}{6N}]$. We may construct the Igusa tower T'_α over $S' = X(N)[\frac{1}{\mathcal{E}}]$ in exactly the same manner as $T_{\alpha/X_1(N)}$, replacing $X_1(N)$ by $X(N)$. Again T_α is a finite quotient of T'_α; so, we only need to prove the irreducibility of T'_α. By Zariski's connectedness theorem ([ALG] III.11.5), the irreducibility of the projective curve $X(N)$ over $\overline{\mathbb{Q}}$ combined with $f_*\mathcal{O}_{X(N)} = \mathbb{Z}[\zeta_N, \frac{1}{6N}]$ for $f : X(N) \to Spec(\mathbb{Z}[\zeta_N, \frac{1}{6N}])$ implies the irreducibility of S_1 over $\overline{\mathbb{F}}_p$ (an algebraic closure of $\mathbb{F}_p$). We extend the scalar to $\overline{\mathbb{F}}_p$ and again write V_0 for $V_{1,0} \otimes_{\mathbb{F}_p} \overline{\mathbb{F}}_p$. Then, V_0 is an integral domain over $\overline{\mathbb{F}}_p$. Since T'_α/S'_1 is étale, T'_α is a disjoint union of irreducible curves over $S'_{1/\overline{\mathbb{F}}_p}$. By the existence of the Tate curve, the cusp ∞ gives a point of T'_α, and there exists a unique irreducible component $C_\alpha = C_{\Gamma(N),\alpha}$ of T'_α containing ∞. The action of $\mathbb{Z}_p^\times$ on the level p-structure permutes irreducible components of T'_α; so, all other irreducible components are isomorphic to C_α. In this way, we get a tower of irreducible curves

$$\cdots \to C_\alpha \to \cdots \to C_0 = S'_1.$$

Since $(\mathbb{Z}/p^\alpha\mathbb{Z})^\times$ acts transitively on points of T'_α over a given closed point of S'_1, C_α/S'_1 is a Galois covering with Galois group

$$H_\alpha = \{\sigma \in (\mathbb{Z}/p^\alpha\mathbb{Z})^\times | \sigma(C_\alpha) \subset C_\alpha\}.$$

Thus $H = \varprojlim_\alpha H_\alpha$ is a subgroup of $\mathbb{Z}_p^\times$. We need to show that $H = \mathbb{Z}_p^\times$.

We fix an embedding $i_p : \overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_p$ for an algebraic closure $\overline{\mathbb{Q}}_p$ of $\mathbb{Q}_p$. Let $\mathbb{Q}_p^{ur}$ be the maximal unramified extension of $\mathbb{Q}_p$ inside $\overline{\mathbb{Q}}_p$, and write W for the p-adic completion of the p-adic integer ring of $\mathbb{Q}_p^{ur}$. We write $\mathcal{W} = i_p^{-1}(W)$; so, $\mathcal{W}$ is a strict Henselization of $\mathbb{Z}_{(p)} \subset \mathbb{Q}$, which is the localization in $\mathbb{Q}$ (not a completion) at p of $\mathbb{Z}$. Let R_0 be the affine ring of $S_{0/\mathcal{W}}$; so, $V_0 = R_0/pR_0$ and $S_0 = Spec(R_0)$ for the $\mathcal{W}$-algebra R_0. The ring $\mathcal{W}[\mu_{p^\infty}]$ is a valuation ring (no longer discrete). We write R_α for the integral closure of $R_0[\mu_{p^\infty}]$ in $X(Np^\alpha)_{/\mathcal{K}}$ for the field of fractions $\mathcal{K}$ of $\mathcal{W}[\mu_{p^\infty}]$. Each element $f \in R_\alpha$ has its q-expansion $f(q) = \sum_{n \gg -\infty} a_n(f)q^n$. We then define a valuation $v = v_\alpha : R_\alpha \to \mathbb{Q}$ by $v_\alpha(f) = \mathrm{Inf}_n\, ord_p(i_p(a_n(f)))$, where ord_p is the valuation of $\mathcal{W}$ normalized so that $ord_p(p) = 1$. Obviously, by the same definition, this valuation extends to $v : \mathfrak{K} \to \mathbb{Q} \sqcup \{\infty\}$ for the modular function field $\mathfrak{K}$ in Theorem 2.40. Note that $R_0/pR_0 = V_0$ and v_α is an extension of the valuation v_0 to R_α. Thus for the ideal $\mathfrak{p}_\alpha = \{f \in R_\alpha | v_\alpha(f) > 0\}$ of R_α, we have an embedding $R_\alpha/\mathfrak{p}_\alpha \hookrightarrow \overline{\mathbb{F}}_p(C_\alpha)$ by the q-expansion (irreducibility of C_α). Let κ_α be the field of fractions of $R_\alpha/\mathfrak{p}_\alpha$. Since $\mathbb{F}_p(C_\infty)$ is Galois over $\kappa_0 = \mathbb{F}_p(C_0)$, the

Galois group H surjects down to $\mathrm{Gal}(\kappa_\infty/\kappa_0)$ by the restriction map; so, we need to show that $\mathrm{Aut}(\kappa_\infty/\kappa_0) \supset \mathbb{Z}_p^\times$, where $\kappa_\infty = \bigcup_\alpha \kappa_\alpha$. Let $\mathcal{V}_\alpha$ be the valuation ring of v_α in $\mathcal{K}(X(Np^\alpha))$. By definition, we have $\kappa_\alpha = \mathcal{V}_\alpha/\mathfrak{P}_\alpha$ for the maximal ideal $\mathfrak{P}_\alpha$ of $\mathcal{V}_\alpha$. Note that $\mathfrak{G} = \mathrm{Gal}(\mathfrak{K}/\mathbb{Q}(J)) = GL_2(\widehat{\mathbb{Z}})/\{\pm 1\}$ by Theorem 2.40. By the description of the action of $\tau(\alpha)$ on $\mathfrak{K}$ in Section 2.6, we find that the action of $\left(\begin{smallmatrix} a & 0 \\ 0 & a^{-1} \end{smallmatrix}\right)$ with $a \in \mathbb{Z}_p^\times$ is induced by the action $\phi_p \mapsto \phi_p \circ \left(\begin{smallmatrix} a & 0 \\ 0 & a^{-1} \end{smallmatrix}\right)$ on level $\Gamma(p^\infty)$-structures ϕ_p. By our choice, C_α contains the infinity cusp where the Tate curve $\mathrm{Tate}(q)_{/\mathbb{Z}[[q^{1/p^\alpha}]]}$ has the level structure $\phi_{can} : \mu_{p^\alpha} \times \mathbb{Z}/p^\alpha\mathbb{Z} \to \mathrm{Tate}(q)[p^\alpha]$ given by $(\zeta, m) \mapsto (\zeta q^{m/p^\alpha}) \in \mathbb{G}_m(W[[q^{1/p^\alpha}]])/q^{\mathbb{Z}}$. Thus the action of a diagonal matrix $\left(\begin{smallmatrix} a & 0 \\ 0 & a^{-1} \end{smallmatrix}\right)$ brings the étale part $\phi^{et}_{can} = \phi_{can}|_{\mathbb{Z}/p^\alpha\mathbb{Z}}$ to $\phi^{et}_{can} \circ a$ for $a \in \mathbb{Z}_p^\times$, which induces the action of $a \in \mathbb{Z}_p^\times$ on $V_\infty = \bigcup_\alpha V_\alpha$. Hence we need to prove that $D/I \supset \mathbb{Z}_p^\times$ by $a \mapsto \left(\begin{smallmatrix} a & 0 \\ 0 & a^{-1} \end{smallmatrix}\right)$ for the decomposition group $D \subset \mathfrak{G}$ and the inertia group $I \subset \mathfrak{G}$ of $\mathfrak{P}_\infty/\mathfrak{P}_0$. In other words, we need to show that for a diagonal matrix $\alpha \in SL_2(\mathbb{Q}_p)$, $\tau(\alpha)$ preserves the valuation $v : \mathfrak{K} \to \mathbb{Q} \sqcup \{\infty\}$.

Let D_p be the decomposition subgroup in $\mathrm{Aut}(\mathfrak{K})$ of v; that is,

$$D_p = \left\{\sigma \in \mathrm{Aut}(\mathfrak{K}) \middle| v \circ \sigma = v\right\}.$$

We note here that D_p is a closed subgroup of $\mathrm{Aut}(\mathfrak{K})$. Indeed, by Hilbert's theory of decomposition groups (e.g. [BCM] V.2.3), for any subfield $\mathfrak{k} \subset \mathfrak{K}$ over which $\mathfrak{K}$ is algebraic, the intersection $D_\mathfrak{k} = D_p \cap \mathrm{Aut}(\mathfrak{K}/\mathfrak{k})$ is compact by the surjectivity of $D_\mathfrak{k}$ onto $\mathrm{Aut}(\kappa_\mathfrak{K}/\kappa_\mathfrak{k})$ for the residue fields $\kappa_\mathfrak{K}$ and $\kappa_\mathfrak{k}$ of the valuation v restricted to $\mathfrak{K}$ and $\mathfrak{k}$, respectively. Note here $\{\mathrm{Gal}(\mathfrak{K}/\mathfrak{k})\}_\mathfrak{k}$ for $\mathfrak{k}$ finitely generated over $\mathbb{Q}$ with algebraic $\mathfrak{K}/\mathfrak{k}$ gives a system of open neighborhoods of the identity of $\mathrm{Aut}(\mathfrak{K})$; so, D_p is a closed subgroup (see Section 2.3 for the topology of the automorphism group of a field).

As we have seen in Section 2.6, the action of a rational upper triangular matrix $\alpha = \left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right) \in GL_2(\mathbb{Q})_+$ on $f(q)$ is given by $f \mapsto f^{\tau(\alpha)} = f \circ \alpha$ which sends $\sum_m a_m(f)q^m$ to $\sum_m \exp(2\pi i \frac{mb}{d}) a_m(f) q^{ma/d}$. Thus $\tau(\alpha)$ preserves v. In other words, $D_p \supset B(\mathbb{Q}) = \left\{\left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right) \in GL_2(\mathbb{Q})_+ \middle| a, b, d \in \mathbb{Q}\right\}$. Since $U(\mathbb{Q}) \cong \mathbb{Q}$ is dense in $U(\mathbb{A}^{(\infty)}) \cong \mathbb{A}^{(\infty)}$ for the upper unipotent radical U, we find $U(\mathbb{A}^{(\infty)}) \overset{\tau}{\subset} D_p$.

Since $SL_2(\mathbb{A}^{(p\infty)})$ in $\mathrm{Aut}(\mathfrak{K})$ commutes with the diagonal torus $T(\mathbb{Z}_p)$ given by $\left\{\left(\begin{smallmatrix} a & 0 \\ 0 & a^{-1} \end{smallmatrix}\right) \middle| a \in \mathbb{Z}_p^\times\right\}$, the action of $g \in SL_2(\mathbb{A}^{(p\infty)})$ can be lifted to an action of the function field $\bigcup_{N,\alpha} \overline{\mathbb{F}}_p(C_{\Gamma(N),\alpha})$; so, v is preserved by $SL_2(\mathbb{A}^{(p\infty)})$ (this point is addressed more carefully in the proof of Theorem 4.21). Since the closure of $SL_2(\mathbb{A}^{(p\infty)})B(\mathbb{Q})$ in $GL_2(\mathbb{A})_+/\mathbb{Q}^\times\mathbb{R}_+^\times = \mathrm{Aut}(\mathfrak{K})$ (see Theorem 2.40) is contained in the closed subgroup D_p and contains the diagonal torus $T(\mathbb{Z}_p)$, we conclude that $T(\mathbb{Z}_p)$ preserves v. □

The following corollary is called the q-expansion principle and is a consequence of the irreducibility of T_α combined with the fact that the local q-parameter gives the uniformizing parameter at a cusp of $X_1(N)$ over $\mathbb{Z}[\frac{1}{N}]$.

Corollary 3.4 *Let $\mathcal{L}_{/\overline{\mathbb{F}}_p}$ be a line bundle realized in the function field of $T_{\alpha/\overline{\mathbb{F}}_p}$. Let s be a cusp and U be a Zariski open neighborhood of $s \in T_{\alpha/\overline{\mathbb{F}}_p}$. If the q-expansion at s of a section $f \in H^0(U, \mathcal{L})$ vanishes, then f vanishes on U.*

3.2.9 p-Adic Elliptic Modular Forms

Since S_m is affine, $T_{m,\alpha}$ is also affine. We write $T_{m,\alpha} = Spec(V_{m,\alpha})$. We have a tower of W_m-flat algebras:

$$V_{m,0} \subset V_{m,1} \subset \cdots \subset V_{m,\alpha} \subset \cdots .$$

These algebras are unramified over $V_{m,0}$ and $\mathrm{Gal}(V_{m,\alpha}/V_{m,0}) \cong (\mathbb{Z}/p^\alpha\mathbb{Z})^\times$. By construction, we have $T_{m,\infty} = Spec(V_{m,\infty})$ for $V_{m,\infty} = \bigcup_\alpha V_{m,\alpha}$. Write simply P_α for $\mathbf{E}[p^\alpha]^{et}$ as a locally free étale group scheme over S_m. Over $V_{m,\alpha}$, we have a canonical isomorphism (the universal level p-structure)

$$I_{can} = \boldsymbol{\phi}_{p^\alpha} : \mathbb{Z}/p^\alpha\mathbb{Z} \cong P_\alpha = \mathbf{E}[p^\alpha]^{et}.$$

We then define $V_{m,\infty} = \bigcup_\alpha V_{m,\alpha}$ and

$$\mathcal{V} = \mathcal{V}_{\Gamma_1(N)} = \varinjlim_m V_{m,\infty} \quad \text{and} \quad V = V_{\Gamma_1(N)} = \varprojlim_m V_{m,\infty}.$$

The space $V_{\Gamma_1(N)}$ is the space of p-adic modular forms on $\Gamma_1(N)$. By taking the Cartier dual of $\mathbb{Z}/p^\alpha\mathbb{Z} \hookrightarrow E[p^\alpha]$, we may regard $f \in V_{m,\alpha}$ as a function of $(E, \phi_p : \mu_{p^\alpha} \hookrightarrow E[p^\alpha], \phi_N)$ satisfying the following conditions similar to (G1–3) for p-adically complete W-algebras $R = \varprojlim_n R/p^n R$.

(G_p1) $(E, \phi_p, \phi_N)_{/R} \cong (E', \phi'_p, \phi'_N)_{/R} \Rightarrow f(E, \phi_p, \phi_N) = f(E', \phi'_p, \phi'_N) \in R$;
(G_p2) If $\rho : R \to R'$ is a continuous W-algebra homomorphism, then

$$f((E, \phi_p, \phi_N) \times_{R,\rho} R') = \rho(f(E, \phi_p, \phi_N));$$

(G_p3) For all level N-structures ϕ_N of type $\Gamma_1(N)$ on the Tate curve $\mathrm{Tate}(q)$,

$$f(\mathrm{Tate}(q), \phi_p^{can}, \phi_N) \in W[[q^{1/N}]].$$

Each element $f \in V_{\Gamma_1(N)} \widehat{\otimes}_W \mathcal{B}$ for a W-algebra $\mathcal{B} = \varprojlim_m \mathcal{B}/p^m\mathcal{B}$ is a function of $(E, \phi_p, \phi_N)_{/R}$ satisfying the conditions (G_p1–3) for p-adically complete $\mathcal{B}$-algebras R (replacing W in the conditions by $\mathcal{B}$).

Again we note the consequence of the irreducibility of the Igusa curve.

Corollary 3.5 (q-Expansion Principle) *If the q-expansion at the infinity cusp of a p-adic modular form f over a p-adically complete W-algebra $\mathcal{B}$ vanishes, f itself vanishes; that is, $f(E, \phi_p, \phi_N) = 0$ for all test objects $(E, \phi_p, \phi_N)_{/R}$ defined over all p-adically complete $\mathcal{B}$-algebras R.*

3.3 p-Ordinary Elliptic Modular Forms

We have a p-divisible module $\mathcal{V}_{\Gamma_1(N)}$ on which $\mathrm{Gal}(V_{m,\infty}/V_{m,0}) = \mathbb{Z}_p^\times = T(\mathbb{Z}_p)$ acts continuously. Here $T = \mathbb{G}_m$. We construct a projector e acting on $\mathcal{V}$ out of the Hecke operator $U(p)$ commuting with the action of $\mathbb{Z}_p^\times = \mathrm{Gal}(V_{m,\infty}/V_{m,0})$. The important features of e are

- $e = \lim_{n\to\infty} U(p)^{n!}$ (we recall the definition of $U(p)$ in Section 3.3.3);
- $\mathcal{V}^{ord} = e\mathcal{V}$ has the Pontryagin dual which is projective over $W[[\mathbb{G}_m(\mathbb{Z}_p)]]$;
- For any $k \geq 3$, there is a canonical isomorphism

$$\mathcal{V}^{ord}[k] \cong eH^0(S_0, \underline{\omega}^k \otimes \mathbb{T}_p) = eH^0(M, \underline{\omega}^k \otimes \mathbb{T}_p) \quad (\mathbb{T}_p = \mathbb{Q}_p/\mathbb{Z}_p),$$

where $\mathcal{V}[k] = \{f \in \mathcal{V} | zf = z^k f \ \forall z \in \mathbb{Z}_p^\times\}$.

Here is how to prove the convergence of the projector $e = \lim_{n\to\infty} U(p)^{n!}$. For any finite p-torsion module X with an endomorphism u, the power $u^{n!}$ stabilizes as n grows, giving an idempotent e_u in the endomorphism algebra $\mathrm{End}(X)$, since $\mathrm{End}(X)$ is a finite ring. For a p-divisible module X with an endomorphism u, we assume that $X = \varinjlim_{\imath\in I} X_\imath$ for finite modules $X_\imath$ stable under u. Then the endomorphism $e_u = \lim_{n\to\infty} u^{n!}$ is well-defined over $X_\imath$ for all i and hence over X. For any compact p-profinite module X with a continuous endomorphism u, we take the Pontryagin dual pair (X^*, u^*) which is p-divisible (cf. [FAN] Chapters 1 and 3). Thus e_{u^*} exists on X^*. Then $e_u = \lim_{n\to\infty} u^{n!}$ (which is the dual of e_{u^*}) exists as an endomorphism of X.

3.3.1 Axiomatic Treatment

Let $\underline{\omega}^k = \mathcal{O}(k)$ for the embedding of $X_1(N) = \widetilde{Proj(G_{\Gamma_1(N)})}$ into the projective space. Thus $\underline{\omega}^k$ is the quasi-coherent sheaf $\widetilde{G_{\Gamma_1(N)}(k)}$ associated with the graded piece $G_{\Gamma_1(N)}(k)$ ([GME] 1.3.3). Then $\underline{\omega}^k = \underline{\omega}^{\otimes k}$. Computing the genus of $X_1(N)$, the Riemann–Roch theorem tells us that $\underline{\omega}^k$ is very ample if $k \geq 3$ (see [GME] Proposition 2.1.4 and Theorem 3.1.2). Therefore $\underline{\omega}^k$ is the pull back of $\mathcal{O}(k)$ of the target projective space.

Let $(\mathbf{E}, \boldsymbol{\phi}_N, \boldsymbol{\omega})$ be the universal elliptic curve over $Y_1(N)$. For each triple (E, ϕ_N, ω) defined over R (called a *test object*), we have a unique morphism $\iota : Spec(R) \to Y_1(N)$ such that $\iota^*(\mathbf{E}, \boldsymbol{\phi}_N, \boldsymbol{\omega}) = (E, \phi_N, \omega)$. For each section $f \in H^0(Y_1(N), \underline{\omega}^k)$, we define $\iota^* f = f(E, \phi_N, \omega)\omega^{\otimes k}$. The function $(E, \phi_N, \omega) \mapsto f(E, \phi_N, \omega)$ satisfies (G0–2). The condition (G3) ensures that f extends to $X_1(N)$. This shows

$$H^0(X_1(N)_{/R}, \underline{\omega}^k) \cong G_k(\Gamma_1(N); R) \quad \text{(canonically)} \tag{3.1}$$

for all $\mathbb{Z}[\frac{1}{6N}]$-algebras R.

Let $(\mathbf{E}, \boldsymbol{\phi}_p, \boldsymbol{\phi}_N)$ be the universal elliptic curve over S_m. Pick a section $f \in H^0(S_m, \underline{\omega}^k)$. Since μ_{p^∞} carries a canonical differential $\omega_{can} = \frac{dt}{t}$, writing

$\mu_{p^\alpha} = Spec(\mathbb{Z}[t]/(t^{p^\alpha}-1))$, we may regard f as a function of (E, ϕ_p, ϕ_N) by $f(E, \phi_p, \phi_N) = f(E, \phi_N, \phi_{p,*}\omega_{can})$. For each $(E, \phi_p, \phi_N) \in \mathcal{E}^{ord}_\infty(R)$ for a W_m-algebra R, we have a unique morphism $\iota : Spec(R) \to T_{m,\infty} = Spec(V_{m,\infty})$ such that $(E, \phi_p, \phi_N) = \iota^*(\mathbf{E}, \boldsymbol{\phi}_p, \boldsymbol{\phi}_N)$. Then $\iota^* f$ is just a function on $f(E, \phi_p, \phi_N)$ such that $f(E, z^{-1}\phi_p, \phi_N) = z^k f(E, \phi_p, \phi_N)$ for $z \in \mathrm{Gal}(V_{m,\infty}/V_{m,0}) = \mathbb{Z}_p^\times$. We recall that the action of $T(\mathbb{Z}_p) = \mathbb{Z}_p^\times$ is normalized so that $z \in \mathbb{Z}_p^\times$ acts on the dual identification $\phi_p^{et} : E[p^\infty]^{et} \cong \mathbb{Q}_p/\mathbb{Z}_p$ of $\phi_p : \mu_{p^\infty} \hookrightarrow E[p^\infty]$ via the multiplication by z; so, z acts on ϕ_p by $\phi_p \mapsto \phi_p \circ z^{-1}$. In other words, we have $z \cdot f(E, \phi_p, \phi_N) = f(E, z^{-1}\phi_p, \phi_N) = z^k f(E, \phi_p, \phi_N)$. This shows that

$$V_{m,\infty}[k] = H^0(S_m, \underline{\omega}^k)$$
$$\text{and} \quad \mathcal{V}[k] = H^0(S_{\infty/W}, \underline{\omega}^k \otimes \mathbb{T}_p) = H^0(S_{\infty/W}, \underline{\omega}^k) \otimes \mathbb{T}_p, \quad (3.2)$$

where $\mathbb{T}_p = \mathbb{Q}_p/\mathbb{Z}_p$. The last identity follows from the affineness of S_0. When we generalize this point to classical groups bigger than $SL(2)_{/\mathbb{Q}}$, the affineness is no longer true; so, we need to work out this point (the base change property) carefully: how to bring $\mathbb{T}_p$ inside the cohomology group. In the elliptic modular case, simply by the affineness of S_0, (3.2) shows that $\mathcal{V}[k]$ is p-divisible, and its direct summand $e\mathcal{V}[k]$ is also p-divisible.

We consider the following condition:

(F) $\mathrm{corank}_W e\mathcal{V}[k] = \mathrm{rank}_W \mathrm{Hom}(e\mathcal{V}[k], \mathbb{T}_p)$ *is finite for an integer* k.

In practice, this condition is often proven by showing

$$eH^0(S_{0/W}, \underline{\omega}^k \otimes \mathbb{T}_p) = eH^0(X_1(N)_{/W}, \underline{\omega}^k \otimes \mathbb{T}_p) = G_k^{ord}(\Gamma_1(N); \mathbb{T}_p), \quad (3.3)$$

where

$$G_k^{ord}(\Gamma_1(N); R) = eG_k(\Gamma_1(N); R) \text{ and } G_k(\Gamma_1(N); \mathbb{T}_p) = G_k(\Gamma_1(N); \mathbb{Z}_p) \otimes \mathbb{T}_p.$$

The left-hand side (LHS) of (3.3) is p-divisible, since S_0 is affine. The right-hand side (RHS) is of finite corank since $X_1(N)$ is projective. Thus $e\mathcal{V}[k]$ is p-divisible of finite corank.

Decompose $\mathbb{Z}_p^\times = \Gamma_T \times \Delta$ for a p-profinite group Γ_T and a prime-to-p finite group Δ. For simplicity suppose that $p > 2$. Then Γ_T is isomorphic to $\mathbb{Z}_p$ and for its generator γ, we have $W[[\Gamma_T]] \cong W[[X]] = \Lambda$ via $\gamma \mapsto 1 + X$ (i.e., $\gamma^s \mapsto (1+X)^s = \sum_{\jmath=0}^\infty \binom{s}{\jmath} X^\jmath$), and $W[[\mathbb{Z}_p^\times]] = \Lambda[\Delta]$. Let V^{ord} be the Pontryagin dual module of $e\mathcal{V}$.

See [FAN] Chapters 1 and 3 for generality on topological groups (including Pontryagin duality and profinite groups). My books, [MFG] Chapter 2 and [LFE] Chapter 8, also have some short exposition on this subject. The Pontryagin dual for p-torsion modules X (equipped with discrete topology) is given by $X^* = \mathrm{Hom}_{cont}(X, \mathbb{T}_p)$ for $\mathbb{T}_p = \mathbb{Q}_p/\mathbb{Z}_p$ with discrete topology. Writing $X = \varinjlim_{\imath \in I} X_\imath$ for finite submodules $X_\imath$, we have

$X^* = \varprojlim_{\iota,n} \mathrm{Hom}(X_\iota, p^{-n}\mathbb{Z}/\mathbb{Z})$, and the topology on X^* is the profinite topology with respect to the finite discrete groups $\mathrm{Hom}(X_\iota, p^{-n}\mathbb{Z}/\mathbb{Z})$. Then X^* is a compact profinite group. The duality is perfect; so, if $X \hookrightarrow Y \twoheadrightarrow Z$ is an exact sequence of p-torsion modules, $Z^* \hookrightarrow Y^* \twoheadrightarrow X^*$ is an exact sequence of profinite groups. In particular, if X is p-divisible and p-torsion, the multiplication by p is surjective; so, by duality, multiplication by p on X^* is injective, which shows that X^* is torsion-free if X is p-divisible and p-torsion.

Suppose that a profinite (Noetherian) W-algebra R acts on X continuously. For any closed ideal $\mathfrak{a} \subset R$, Pontryagin's dual of $X[\mathfrak{a}] = \{x \in X | \mathfrak{a}x = 0\}$ is given by $X^*/\mathfrak{a}X^*$, where R acts on X^* continuously by the pullback (dual) action. To see this, we choose a finite set of generators $(a_1, \ldots, a_r)$ of $\mathfrak{a}$ and look into the exact sequence: $0 \to X[\mathfrak{a}] \to X \xrightarrow{x \mapsto (a_1 x, \ldots, a_r x)} X^r$. Taking the dual, we find that $X^{*r} \xrightarrow{(x_1,\ldots,x_r) \mapsto \sum_\iota a_\iota x_\iota} X^* \to X^*/\mathfrak{a}X^* \to 0$ is exact, which shows $X[\mathfrak{a}]^* = X^*/\mathfrak{a}X^*$.

For any (continuous) W-algebra homomorphism $\chi : R \to W$, applying the above fact to $\mathfrak{a} = \mathrm{Ker}(\chi)$, the Pontryagin dual of the χ-eigen subspace

$$X[\chi] = \{x \in X | ax = \chi(a)x \ \text{ for all } a \in R\}$$

is given by the maximal χ-eigen quotient $X^*/\mathfrak{a}X^*$ for the χ-augmentation ideal $\mathfrak{a} = \mathrm{Ker}(\chi) = (a - \chi(a))_{a \in R} \subset R$.

Let $\widehat{\Delta} = \mathrm{Hom}(\Delta, W^\times)$, and write $\mathcal{V}[\chi]$ for the χ-eigen subspace of $\chi \in \widehat{\Delta}$:

$$\mathcal{V}[\chi] = \{v \in \mathcal{V} | \delta v = \chi(\delta)v \ \text{ for all } \delta \in \Delta\}.$$

If $e\mathcal{V}[k]$ is of finite corank for one k, taking $\chi \in \widehat{\Delta}$ with $\chi = k|_\Delta$, we have $\mathcal{V}[k] = \mathcal{V}[\chi][k]$, and then by duality, we have

$$\begin{aligned} V^{ord} \otimes_{W[[T(\mathbb{Z}_p)]],k} W &= V^{ord}[\chi]/(X + 1 - \gamma^k)V^{ord}[\chi] \\ &= \mathrm{Hom}_{cont}(\mathcal{V}^{ord}[\chi, (X + 1 - \gamma^k)], \mathbb{T}_p) \\ &\cong \mathrm{Hom}_{\mathbb{Z}_p}(\varinjlim_m H^0_{ord}(X_1(N)_{/W_m}, \underline{\omega}^k_{/W_m}), \varinjlim_m p^{-m}\mathbb{Z}/\mathbb{Z}) \\ &\cong \varprojlim_m \mathrm{Hom}_{\mathbb{Z}_p}(H^0_{ord}(X_1(N)_{/W_m}, \underline{\omega}^k_{/W_m}), p^{-m}\mathbb{Z}/\mathbb{Z}) \\ &\cong \varprojlim_m \mathrm{Hom}_W(H^0_{ord}(X_1(N)_{/W_m}, \underline{\omega}^k_{/W_m}), W_m) \\ &= \mathrm{Hom}_W(H^0_{ord}(X_1(N), \underline{\omega}^k), W) \cong \mathrm{Hom}_W(G^{ord}_k(\Gamma_1(N); W), W). \end{aligned} \tag{3.4}$$

In the middle equality, we have assumed (3.3). Here the subscript or superscript *ord* indicates the image of e.

Write $\widehat{\Delta} = \mathrm{Hom}(\Delta, W^\times)$. Note here that each element $\chi \in \widehat{\Delta}$ is induced by (infinitely many) positive $k \in X(T)$. Decompose V^{ord} by the character of Δ as follows,

$$V^{ord} = \bigoplus_{\chi \in \widehat{\Delta}} V^{ord}[\chi].$$

This is possible as we assumed that Δ has order prime to p. If $z \mapsto z^k$ coincides with χ on Δ, then $V^{ord}[\chi] \otimes_{\Lambda,k} W = V^{ord} \otimes_{W[[T(\mathbb{Z}_p)]],k} W$. By Nakayama's lemma, we have a surjective homomorphism of Λ-modules:

$$\pi : \Lambda^{s(\chi)} \twoheadrightarrow V^{ord}[\chi],$$

where $s = s(\chi) = \operatorname{corank}_W e\mathcal{V}[k]$. If (F) holds for one k, it holds for all $\kappa \in X(T)$ inducing χ, and π has to be an isomorphism for the following reason: The number s is the minimum number of generators of $V^{ord}[\chi] \otimes_{\Lambda,\kappa} W$ over W. We know that this module is W-free, because its dual $\mathcal{V}[\kappa]$ is p-divisible; so, it is free of rank s. The morphism π induces an isomorphism modulo $(1+X) - \gamma^\kappa$ for all κ inducing χ. Then $\operatorname{Ker}(\pi) \subset \bigcap_\kappa \operatorname{Ker}(\pi \mod (1 + X - \gamma^\kappa)) = 0$, and we get

Theorem 3.6 *Suppose that* (F) *holds for one k. Then V^{ord} is Λ-free of finite rank, and if* (3.3) *holds for $\kappa \in X(T)$ with $\kappa|_\Delta = k|_\Delta$, then*

$$V^{ord} \otimes_{W[[\mathbb{Z}_p^\times]],\kappa} W \cong \operatorname{Hom}_W(G_\kappa^{ord}(\Gamma_1(N); W), W).$$

3.3.2 Bounding the p-Ordinary Rank

Since S_0 is affine, we have $H^0(S_1, \underline{\omega}^k) = H^0(S_{0/W}, \underline{\omega}^k) \otimes_W W_1$. If we have a sequence $\overline{f}_1, \overline{f}_2, \ldots, \overline{f}_j$ of linearly independent sections in $H^0_{ord}(S_1, \underline{\omega}^k)$, we can lift them to $f_i \in H^0(S_{0/W}, \underline{\omega}^k)$ so that $\overline{f}_i = (f_i \mod p)$. Since $S_0 = M[\frac{1}{\mathcal{E}}]$, we have

$$H^0(S_{0/W}, \underline{\omega}^k) = \varinjlim_n H^0(M_{/W}, \underline{\omega}^{k+n(p-1)})/\mathcal{E}^n.$$

Thus $\mathcal{E}^m f_i \in H^0(M_{/W}, \underline{\omega}^{k+m(p-1)})$ for all $i = 1, \ldots, j$ for sufficiently large m, and they are linearly independent. We now assume

(C) $\quad e(\mathcal{E}f) = \mathcal{E}(ef)$ for all $f \in H^0(S_1, \underline{\omega}^k)$.

By this, $e(\mathcal{E}^m f_i)$ are still linearly independent in $H^0_{ord}(M_{/K}, \underline{\omega}^{k+m(p-1)})$; so, we have $\dim_K H^0_{ord}(M_{/K}, \underline{\omega}^{k+m(p-1)}) \geq j$. If $\operatorname{rank}_{W_1} H^0_{ord}(S_1, \underline{\omega}^k)$ is infinite, the rank of $H^0_{ord}(M_{/K}, \underline{\omega}^{k+m(p-1)})$ grows as m grows. Thus the condition (F) for all k follows from

(F′) $\quad \dim_K G_k^{ord}(\Gamma_1(N); \mathbb{Q}_p)$ is bounded (independently of k).

Indeed, the Eichler–Shimura isomorphism combined with a calculation of group cohomology $H^1_{ord}(\Gamma_1(N), L(k; K))$ proves much stronger

(E) $\quad$ If $k \geq 3$, $\dim_K G_k^{ord}(\Gamma_1(N); K)$ depends only on $k \mod p - 1$

for the projector e. Actually we can ease the condition $k \geq 3$ to $k \geq 2$ if we replace $G_k^{ord}(\Gamma_1(N); K)$ by $G_k^{ord}(\Gamma_1(N) \cap \Gamma_0(p); K)$ in (E) (see Section 4.3 in the text and [LFE] 7.2).

3.3.3 *p*-Ordinary Projector

Let $(E, \phi_p, \phi_N)_{/R}$ be a test object. Each subgroup C of order p outside the image of ϕ_p is étale; so, we can think of the quotient $(E/C, \phi_p, \phi_N)$ defined over a (possibly radiciel) finite extension R' of R. We define the $U(p)$ operator acting on a (p-adic or classical) geometric modular form f by

$$f|U(p)(E, \phi_p, \phi_N) = \frac{1}{p}\sum_C f(E/C, \phi_p, \phi_N).$$

The sum $\sum_C$ in the above definition has to be understood as a trace from R' to R (see (4.18)) if R is not a $\mathbb{Q}$-algebra (this subtle point is discussed in detail later in Section 8.3.1). From the definition, as long as the prime p is invertible in the base ring R, we can verify the conditions (G0–3) (or (G_p1–3)) for $f|U(p)$. Computing the q-expansion, we know $a(n, f|U(p)) = a(np, f)$. Then by the q-expansion principle, the operator $U(p)$ is well-defined p-integrally over $\mathcal{W}$ even if the definition of the operator involves division by the prime p.

Since $\mathcal{E} \equiv 1 \mod p$, we confirm (C). Let

$$G_{\Gamma_1(N)}(R) = \bigoplus_{k=0}^{\infty} G_k(\Gamma_1(N); R).$$

One can prove the p-adic density of $D = G_{\Gamma_1(N)}(W)[\frac{1}{p}] \cap V$ in V (see Corollary 8.4 in the text and [GME] Corollary 3.2.4). Using this fact, we can show that we can write $\mathcal{V}$ as a union of finite modules stable under $U(p)$; so, $e = \lim_{n\to\infty} U(p)^{n!}$ exists. The final result is as follows.

Theorem 3.7 (VCT) *We have*

$$V^{ord} \otimes_{W[[\mathbb{Z}_p^\times]],k} W \cong \begin{cases} \mathrm{Hom}_W(G_k^{ord}(\Gamma_1(N); W), W) & \text{if } k \geq 3, \\ \mathrm{Hom}_W(G_2^{ord}(\Gamma_1(N) \cap \Gamma_0(p); W), W) & \text{if } k = 2. \end{cases}$$

Similarly, if we write $\mathcal{V}^{ord}_{cusp}$ for the subspace of cusp forms in $\mathcal{V}^{ord}$ and write V^{ord}_{cusp} for its Pontryagin dual (i.e., the cuspidal quotient of V^{ord}), the above result is still valid replacing V^{ord} and G_k^{ord} by V^{ord}_{cusp} and S_k^{ord}.

3.3.4 Families of *p*-Ordinary Modular Forms

Recall the decomposition $T(\mathbb{Z}_p) = \mathbb{Z}_p^\times = \Gamma_T \times \Delta$ for a finite group Δ and $\Gamma_T \cong \mathbb{Z}_p$. We have chosen a generator $\gamma \in \Gamma_T$ so that $\mathbb{Z}_p \ni s \mapsto \gamma^s \in \Gamma_T$ induces the isomorphism $\mathbb{Z}_p \cong \Gamma_T$ and $\Lambda = W[[\Gamma_T]] \cong W[[X]]$ by $\gamma \leftrightarrow 1 + X$.

Let $a(n) : \mathcal{V}^{ord} \to \mathbb{T}_p$ be the linear map associating with f its coefficient of q^n in the q-expansion; so, $a(n)$ is in the dual V^{ord}. We now consider

$$G(\chi; \Lambda) = \mathrm{Hom}_\Lambda(V^{ord}[\chi], \Lambda).$$

With each $\Phi \in G(\chi; \Lambda)$, we associate its q-expansion

$$\Phi(q) = \sum_{n=0}^{\infty} a(n; \Phi) q^n \in \Lambda[[q]] \quad \text{with } a(n, \Phi) = \Phi(a(n)) \in \Lambda.$$

Theorem 3.8 *For each $k \geq 2$, we have*

(1) *$G(\chi; \Lambda)$ is Λ-free of finite rank;*
(2) *$G(\chi; \Lambda) \otimes_{\Lambda,k} W \cong G_k^{ord}(\Gamma_1(N) \cap \Gamma_0(p), \chi\omega^{-k}; W)$, where ω is the Teichmüller character modulo p;*
(3) *the above identification is induced by $\Phi \mapsto \sum_n a(n, \Phi)(\gamma^k - 1)q^n \in W[[q]]$.*

Proof. The Λ-freeness follows from the freeness of $V^{ord}[\chi]$. We only prove the assertion when k induces χ on Δ; so, $\chi\omega^{-k}$ is trivial. We have

$$\begin{aligned} G(\chi; \Lambda) \otimes_{\Lambda,k} W &\cong \mathrm{Hom}_{\Lambda}(V^{ord}[\chi], \Lambda) \otimes_{\Lambda,k} W \\ &\cong \mathrm{Hom}_W(V^{ord}[\chi] \otimes_{\Lambda,k} W, W) \cong \mathrm{Hom}_W(\mathrm{Hom}_W(G_k^{ord}(\Gamma_1(N); W), W)) \\ &= G_k^{ord}(\Gamma_1(N); W) \cong G_k^{ord}(\Gamma_1(N) \cap \Gamma_0(p); W). \end{aligned}$$

Here is the explanation of how to obtain the last isomorphism in the above sequence of isomorphisms. As is well known (see (4.59)), the Hecke operator $T(p)$ acts on $f \in G_k(\Gamma_1(N); W)$ by

$$a(n, f|T(p)) = a(pn, f) + p^{k-1} a(\frac{n}{p}, f),$$

where we put $a(\xi, f) = 0$ if $\xi \notin \mathbb{Z}$. Thus in $W[[q]]$, $T(p) \equiv U(p) \mod p$ if $k \geq 2$, and the natural inclusion $G_k(\Gamma_1(N); W) \hookrightarrow G_k(\Gamma_1(N) \cap \Gamma_0(p); W)$ induces an inclusion:

$$G_k^{ord}(\Gamma_1(N); W) \hookrightarrow G_k^{ord}(\Gamma_1(N) \cap \Gamma_0(p); W).$$

By the theory of primitive forms ([MFM] Theorem 4.6.17; see also [H85] Section 3), this inclusion is surjective if $k \geq 3$. We leave the verification of assertion (3) to the reader. □

For a given element $\Phi \in G(\chi; \Lambda)$, we get a family of p-ordinary elliptic modular forms $\{\Phi(\gamma^k - 1)\}_{k \geq 2}$ whose q-expansion coefficients depend (p-adic) analytically on the weight $k \in \mathrm{Hom}_{cont}(T(\mathbb{Z}_p), \mathbb{Z}_p^\times)$. Each element $\Phi \in G(\chi; \Lambda)$ is called a p-ordinary *Λ-adic form*, and $\{\Phi(\gamma^k - 1)\}_{k \geq 2}$ is called the *p-adic analytic family* of modular forms associated with a Λ-adic form Φ.

There is a version of this for $\Gamma_0(N)$ and cusp forms. For $k \geq 2$, we have

1. $\mathcal{F}(\Gamma_0(N), \chi; \Lambda)$ is Λ-free of finite rank;
2. $\mathcal{F}(\Gamma_0(N), \chi; \Lambda) \otimes_{\Lambda,k} W \cong \mathcal{F}_k^{ord}(\Gamma_0(Np), \chi\omega^{-k}; W)$;
3. the above identification is induced by $\Phi \mapsto \sum_n a(n, \Phi)(\gamma^k - 1)q^n \in W[[q]]$,

where χ is a character modulo Np and $\mathcal{F} = G$ and S (see [GME] Chapter 3 for a proof).

3.4 Elliptic Λ-Adic Forms and p-Adic L-functions

We describe how to view Λ-adic forms as p-adic modular forms defined over Λ. Once this is done, we can evaluate Λ-adic forms at elliptic curves, which gives us a convenient method of constructing and analyzing p-adic L-functions. Then, we give a short account of the Λ-adic Eisenstein series and examples of Λ-adic L-functions. All the arguments presented here can be generalized to the Hilbert modular case (see Chapter 4), the Siegel–Hilbert modular case, and the quasi-split unitary case (see Chapter 8).

3.4.1 Generality of Λ-Adic Forms

For simplicity, we assume that $p > 2$ and only consider the Λ-adic forms of level p^∞; thus, we have $N = 1$, and $\Gamma_T = 1 + p\mathbb{Z}_p$ is torsion-free. Let $\Lambda = \mathbb{Z}_p[[T]]$. In the previous section, we have introduced the space $G(\chi; \Lambda)$ of p-ordinary Λ-adic forms, which is a free Λ-module of finite rank with

$$G(\chi; \Lambda) \otimes_{\Lambda,k} \mathbb{Z}_p \cong G_k^{ord}(\Gamma_0(p), \chi\omega^{-k}; \mathbb{Z}_p)$$

for all $k \geq 2$. Here $k : \Lambda \to \mathbb{Z}_p$ is the evaluation at $\gamma^k - 1$ of the power series. If we identify Λ with the Iwasawa algebra $\mathbb{Z}_p[[\Gamma_T]]$ by sending the generator γ of Γ_T to $1 + T$, k is induced by the character $\Gamma_T \ni z \mapsto z^k \in \mathbb{Z}_p^\times$.

We write $G(\Lambda)$ for the Λ-module made of formal q-expansions

$$\Phi = \sum_{n \geq 0} a(n; \Phi)(T) q^n \in \Lambda[[q]],$$

such that $\Phi(\gamma^k - 1) \in V[k]$ for infinitely many k. We have $\bigoplus_\chi G(\chi; \Lambda) \subset G(\Lambda)$, where χ runs over (actually even) powers of Teichmüller characters.

We now consider the space of p-adic modular forms $V_{/\Lambda}$ over Λ of level p^∞. In other words, we make a base change $T_{m,n/\mathbb{Z}_p}$ to $T_{m,n/\Lambda} = T_{m,n/\mathbb{Z}_p} \times_{\mathbb{Z}_p} \Lambda$ and consider p-adic modular forms over Λ. The functions in $V_{/\Lambda} = V \widehat{\otimes}_{\mathbb{Z}_p} \Lambda$ classify couples $(E, \phi : \mu_{p^\infty} \hookrightarrow E[p^\infty])_{/R}$ defined over p-adic Λ-algebras R. This $f \in V_{/\Lambda}$ is a functorial rule assigning the value $f(E, \phi) \in R$ for each couple $(E, \phi)_{/R}$ as above.

This space has two Λ-module structures: one coming from the base ring Λ and another coming from the action of $\mathrm{Gal}(T_{m,\infty}/S_m) = \mathbb{Z}_p^\times$ by the diamond operators $\langle z \rangle$. Let $\nu : \Gamma_T \to \Lambda^\times$ be the universal character given by $\nu(z) = [z] \in \Gamma_T$. Then we can define

$$\mathbb{G}(\Lambda) = \left\{ f \in V_{/\Lambda} \middle| f|\langle z \rangle = \nu(z) f \ \forall z \in \Gamma_T \right\}. \tag{3.5}$$

Each $\Phi \in \mathbb{G}(\Lambda)$ has a q-expansion at ∞: $\Phi(T, q) = \sum_{n \geq 0} a(n; \Phi)(T) q^n$. By definition, we have a natural map $V_{/\Lambda} \otimes_{\Lambda,s} \mathbb{Z}_p \to V_{/\mathbb{Z}_p}$ for each $s : \Lambda \to \mathbb{Z}_p$ taking $\Phi(T)$ to $\Phi(\gamma^s - 1)$ for $s \in \mathbb{Z}_p$. Here the tensor product is taken using

a Λ-module structure induced by the diamond operators. Since on $\mathbb{G}(\Lambda)$, the two Λ-module structures coincide, the map is injective by the q-expansion principle. This map brings $\Phi \in \mathbb{G}(\Lambda)$ to a p-adic modular form of weight s. Therefore, Φ is a Λ-adic form.

Conversely, starting from a Λ-adic form Φ, we regard Φ as a bounded measure on Γ_T having values in $V_{/\mathbb{Z}_p}$. Here we use the fact that $\mathbb{Z}_p[[\Gamma_T]]$ is canonically isomorphic to the measure space on Γ_T by $a(T) \mapsto \int x^s da = a(\gamma^s - 1)$ (see [LFE] Sections 3.3–5). Thus, writing $\mathcal{C}(X, Y)$ for the space of continuous functions $X \xrightarrow{f} Y$ for two topological spaces X and Y, we may regard Φ as a $\mathbb{Z}_p$-linear map: $f \mapsto \int f d\Phi$ of $\mathcal{C}(\Gamma_T, \mathbb{Z}_p)$ into $V_{/\mathbb{Z}_p}$. Then for each test object $(E, \phi)_{/R}$ for a p-adic Λ-algebra R, regarding R as a p-adic $\mathbb{Z}_p$-algebra, we can evaluate $\int f d\Phi \in V_{/\mathbb{Z}_p}$ at $(E, \phi)_{/R}$, getting a bounded $\mathbb{Z}_p$-linear form from the space $\mathcal{C}(\Gamma_T, \mathbb{Z}_p)$ into R, which we write $\Phi(E, \phi)(T) \in R \mathbin{\widehat{\otimes}}_{\mathbb{Z}_p} \Lambda = R[[T]]$. Since R is already a Λ-algebra, the Λ-module structure $\Lambda \mathbin{\widehat{\otimes}} R \to R$ given by $\lambda \otimes r = \lambda r$ induces a surjective algebra homomorphism $m : R \mathbin{\widehat{\otimes}}_{\mathbb{Z}_p} \Lambda \twoheadrightarrow R$. We then define $\Phi(E, \phi)$ by $m(\Phi(E, \phi)(T))$. Then the assignment: $(E, \phi) \mapsto \Phi(E, \phi)$ satisfies the axioms (G_p1–3) governing the p-adic modular forms defined over Λ. It is easy to check that this p-adic modular form is in $\mathbb{G}(\Lambda)$ having the same q-expansion at ∞ as Φ. Thus we have found:

Theorem 3.9 *The subspace $\mathbb{G}(\Lambda) \subset V_{/\Lambda}$ is isomorphic to the space $G(\Lambda)$ of all Λ-adic forms via q-expansion at the cusp ∞. In particular, we have*

$$\bigoplus_{\chi} G(\chi; \Lambda) \cong e(\mathbb{G}(\Lambda))$$

for the p-ordinary projector $e : V_{/\Lambda} \twoheadrightarrow V^{ord}_{/\Lambda}$.

Let $(E, \omega)_{/W}$ be an elliptic curve with complex multiplication by an imaginary quadratic field $M = \mathbb{Q}[\sqrt{-D}]$. We suppose that ω is defined over $\mathcal{W} = W \cap \overline{\mathbb{Q}}$ fixing an embedding $i_p : \overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_p$. Suppose that p splits in M and write $p = \mathfrak{p}\bar{\mathfrak{p}}$. Under this assumption, E is p-ordinary. We may assume that W is the p-adic integer ring of the p-adic completion of the maximal unramified extension of $\mathbb{Q}_p$ (so $W = W(\overline{\mathbb{F}}_p)$, the ring of Witt vectors with coefficients in $\overline{\mathbb{F}}_p$; see [BCM] IX.1) and that $E[\bar{\mathfrak{p}}^\infty]$ is the étale part of $E[p^\infty]$ over W. Thus we have $\phi : \mu_{p^\infty} \cong E[\mathfrak{p}^\infty]$. In this way, we can evaluate a given Λ-adic form Φ at (E, ϕ).

Corollary 3.10 *If $\Phi(E, \phi) = 0$ for infinitely many distinct Es with complex multiplication, then $\Phi = 0$. There exist finitely many elliptic curves $(E_i, \phi_i)_{/W}$ such that any given linear form $G(\chi; \Lambda) \to \Lambda$ is a Λ-linear combination of evaluation at (E_i, ϕ_i).*

If $\Phi \in G(\chi; \Lambda)$ and furthermore, if $f = \Phi(\gamma^k - 1) \in G_k(\Gamma_0(p), \chi\omega^{-k}; \mathcal{W})$, then $f(E, \omega) \in \mathcal{W}$. The morphism $\phi : \mu_{p^\infty} \hookrightarrow E$ induces a canonical differential $\omega_{can} = \phi_* \frac{dt}{t}$. Then $\omega = \Omega_p \omega_{can}$, and we have a result of Katz:

$$f(E,\omega)=\frac{f(E,\omega_{can})}{\Omega_p^k}=\frac{\Phi(E,\phi)(\gamma^k-1)}{\Omega_p^k}\in\mathcal{W}\subset\overline{\mathbb{Q}}. \tag{EQ1}$$

We may assume that $E(\mathbb{C})=\mathbb{C}/R$ for the integer ring R of M. Let w be the variable of $\mathbb{C}$. Then dw induces a canonical differential ω_∞ on $E(\mathbb{C})$, $\omega=\Omega_\infty\omega_\infty$, and we get a result of Shimura (actually dating back to a work of Hurwitz [Hz] for elliptic modular forms):

$$f(E,\omega)=\frac{f(E,\omega_\infty)}{\Omega_\infty^k}=\frac{f(E,\omega_{can})}{\Omega_p^k}\in\mathcal{W}\subset\overline{\mathbb{Q}}. \tag{EQ2}$$

The lattice $O=H_1(E,\mathbb{Z})\subset\mathbb{C}$ is generated over O by a single element $\gamma=1$ and $\Omega_\infty=\int_\gamma\omega$, because $\int_\gamma dw=1$.

3.4.2 Some p-Adic L-Functions

In this subsection, we describe very briefly some p-adic L-functions whose construction stems from the theory of Λ-adic forms. More general and thorough discussion can be found in [LFE] Chapters 7 and 10 and [SGL].

For simplicity, we assume that $p>2$ and only consider the Λ-adic Eisenstein series of level p^∞. Let us fix an even power $\chi=\omega^a$ of the Teichmüller character. For simplicity, we choose $a\not\equiv 0 \mod p-1$; so, χ is nontrivial. We consider the Kubota–Leopoldt p-adic L-function $-\frac{1}{2}L_p(1-s,\chi)=a_0(\gamma^s-1)$ $(\gamma=1+p)$ with $a_0\in\mathbb{Z}_p[[T]]$ (e.g., [LFE] Sections 3.5 and 4.4). We have, for positive integers k,

$$L_p(1-k,\chi)=(1-\chi\omega^{-k}(p)p^{k-1})L(1-k,\chi\omega^{-k}).$$

Then we define an element $\mathcal{E}_\chi\in G(\chi;\Lambda)$ by

$$a(n,\mathcal{E}_\chi)=\sum_{0<d|n,p\nmid d}\chi(d)d^{-1}(1+T)^{\log(d)/\log(\gamma)}\quad\text{and}\quad a(0,\mathcal{E}_\chi)=a_0(T).$$

We want to relate $\mathcal{E}_\chi(\gamma^k-1)$ to the following classical Eisenstein series:

$$E_k(E,\omega)=\frac{1}{2}\sum_{(m,n)\neq(0,0)}\frac{1}{(mw_1+nw_2)^k},$$

where $(E,\omega)_{/\mathbb{C}}$ corresponds to the lattice $L=\mathbb{Z}w_1+\mathbb{Z}w_2$ by Weierstrass' theory ([GME] Section 2.4), that is, $E(\mathbb{C})=\mathbb{C}/L$ and $\omega=dw$ for the variable $w\in\mathbb{C}$. As is well known ([IAT] (2.2.1)), for even $k>2$,

$$E_k=c(k)\left(\frac{1}{2}\zeta(1-k)+\sum_{n>0}\Big(\sum_{0<d|n}d^{k-1}\Big)q^n\right)$$

for $c(k)=\frac{\Gamma(k)}{(2\pi\sqrt{-1})^k}$. This shows that if $\omega^k=\chi$, then

$$\mathcal{E}_\chi(\gamma^k - 1) = c(k)^{-1}\left(E_k - p^{k-1}E_k(pz)\right).$$

If we take the elliptic curve (E, ω) defined by $y^2 = 4x^3 - x$ (which is equivalent to $y^2 = 1 - x^4$: Gauss' lemniscate) with $\omega = \frac{dx}{y}$, then it has complex multiplication by $\mathbb{Q}[\sqrt{-1}]$ $(\sqrt{-1} : (x, y) \mapsto (\sqrt{-1}x, y)$ for the model given by $y^2 = 1 - x^4$ and for $k > 2$ with $\omega^k = \chi$, we have

$$\frac{1}{2}E_k(E, \omega) = \frac{L(k, \lambda_k)}{\Omega_\infty^k},$$

where λ_k is a Hecke character of conductor 1 such that $\lambda_k(\alpha) = \alpha^k$. Since we can find $\gamma \in SL_2(\mathbb{Z})\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right) SL_2(\mathbb{Z})$ such that $(w_1, w_2)\gamma$ is a base of $\mathfrak{p}$, we rediscover p-adic interpolation of Hurwitz numbers (done by many outstanding mathematicians):

$$\begin{aligned}\frac{\mathcal{E}_\chi(E, \omega)(\gamma^k - 1)}{\Omega_p^k} =& 2c(k)^{-1}\frac{L(k, \lambda_k) - p^{k-1}\lambda_k(\overline{\mathfrak{p}})p^{-k}L(k, \lambda_k)}{\Omega_\infty^k} \\ =& 2c(k)^{-1}(1 - p^{-1}\lambda_k(\overline{\mathfrak{p}}))\frac{L(k, \lambda_k)}{\Omega_\infty^k}.\end{aligned}$$

This is a Λ-adic version of Katz's way of constructing the p-adic Hecke L-function: $L_p(s) = \mathcal{E}_\chi(E, \omega)(\gamma^s - 1)$ ([K3], [HT], and [SGL] Chapter 8).

A p-adic Rankin product can be constructed similarly. Let Φ be a normalized Hecke eigenform in $G(\chi; \Lambda)$. Write $\mathbb{L}$ for the field of fractions of Λ. As shown in Corollary 5.4, the Hecke algebra acts semi-simply on $G(\chi; \Lambda)$; so, we can uniquely decompose $G(\chi; \Lambda) \otimes_\Lambda \mathbb{L} = \mathbb{L}\Phi \oplus \mathbb{X}$ as Hecke modules. Let $\ell : G(\chi; \Lambda) \to \mathbb{L}$ be the linear form defined by $\Psi = \ell(\Psi)\Phi + x$ for $x \in \mathbb{X}$. We consider two copies of Λ, say $\mathbb{Z}_p[[S]]$ and $\mathbb{Z}_p[[T]]$. Take two Hecke eigenforms $\Phi \in G(\chi; \mathbb{Z}_p[[T]])$ and $\Psi \in G(\psi; \mathbb{Z}_p[[S]])$. Linearly extend ℓ to $G(\chi; \mathbb{Z}_p[[T]]) \widehat{\otimes}_{\mathbb{Z}_p} \mathbb{Z}_p[[S]] \to \mathbb{L} \widehat{\otimes}_{\mathbb{Z}_p} \mathbb{Z}_p[[S]]$. We define

$$L_p(S, T) = \ell(e(\Psi(S)\mathcal{E}_{\chi\psi^{-1}}((1 + T)(1 + S)^{-1} - 1))),$$

where $e : V_{/\mathbb{Z}_p[[T]]} \to V^{ord}_{/\mathbb{Z}_p[[T]]}$ is the p-ordinary projector. Then we see

$$\Psi(S)\mathcal{E}_{\chi\psi^{-1}}(\frac{1 + T}{1 + S} - 1)|_{S=\gamma^m-1, T=\gamma^k-1} = \Psi(\gamma^m - 1)\mathcal{E}_{\chi\psi^{-1}}(\gamma^{k-m} - 1).$$

Thus $\Psi(\gamma^m - 1)\mathcal{E}_{\chi\psi^{-1}}(\gamma^{-m}(1 + T) - 1) \in G(\chi; \mathbb{Z}_p[[T]])$ and hence the p-adic L-value $L_p(\gamma^m - 1, \gamma^k - 1)$ is the coefficient of $\Psi(\gamma^m - 1)E_{k-m}$ in $\Phi(\gamma^k - 1)$ for a suitable Eisenstein series E_{k-m} of weight $k - m$. As is shown by Shimura [Sh7], this coefficient can be computed by the Rankin product value

$$\frac{D(k - 1, \Phi(\gamma^k - 1), \Psi(\gamma^m - 1))}{(\Phi(\gamma^k - 1), \Phi(\gamma^k - 1))} \quad (k > m)$$

for the Petersson inner product $(\ ,\ )$ up to an explicit constant; so, L_p gives p-adic interpolation of the Rankin product. For an explicit evaluation formula for $L_p(S,T)$, see [LFE] Chapters 7 and 10 and [SGL] Chapter 6.

We summarize the basic ingredients we have used in the construction of p-adic L-functions and in the study of Λ-adic forms:

1. The vertical control theorem;
2. The q-expansion principle (irreducibility of the Igusa tower).

We generalize these points to more general automorphic forms in the following chapters.

4

Hilbert Modular Varieties

We present here a general theory of the Hilbert modular varieties together with a sketch of a proof of the geometric reciprocity laws and the vertical control theorems, assuming some of the results which are proven for more general reductive groups G in later chapters.

We start with an affine group scheme H over a ring R'. For a subalgebra R of R', if the covariant functor $C \mapsto H(C \otimes_R R')$ defined on the category of R-algebras is representable by a scheme $H'_{/R}$, we write $H'_{/R} = \mathrm{Res}_{R'/R}H$ and call it the *Weil restriction* of H with respect to R'/R (this is not changing the base ring of $H_{/R'}$ to the subalgebra R). Since $C \mapsto H(C \otimes_R R')$ is a group functor, $\mathrm{Res}_{R'/R}H$ is a group scheme if it exists. If R'/R is locally free of finite rank, an affine group scheme $\mathrm{Res}_{R'/R}H$ always exists (see [NMD] 7.6, Theorem 4).

To connect with what we have said in the introduction (Chapter 1), we describe briefly the algebraic groups G, G_1 and their torus T which gives the notion of weight for Hilbert modular forms. Hilbert modular forms are defined with respect to a fixed base totally real field F with discriminant $d(F)$. We always assume in this section that *the fixed prime p is unramified in $F/\mathbb{Q}$*; so, $p \nmid d(F)$. Write $O = O_F$ for the integer ring of F. Our classical groups are given by $G = \mathrm{Res}_{O/\mathbb{Z}}GL(2)$, $G_1 = \mathrm{Res}_{O/\mathbb{Z}}SL(2)$, and $PG = G^{ad} = \mathrm{Res}_{O/\mathbb{Z}}PGL(2)$. In other words, for commutative algebras R, we define (as functors from algebras into groups) $G(R) = GL_2(R \otimes_{\mathbb{Z}} O)$, $G_1(R) = SL_2(R \otimes_{\mathbb{Z}} O)$, and so on. The torus T is the subgroup made up of diagonal matrices in G_1; so, $T(R) = (R \otimes_{\mathbb{Z}} O)^\times$, and hence $T(\mathbb{Z}_p) = O_p^\times$ for the p-adic completion $O_p = \varprojlim_n O/p^nO$. Writing I for the set of all embeddings of F into $\overline{\mathbb{Q}}$, we may identify $X(T)$ with the formal free module $\mathbb{Z}[I]$ generated by elements of I, associating a character $(x \mapsto x^k = \prod_\sigma \sigma(x)^{k_\sigma})$ with $k = \sum_\sigma k_\sigma \sigma \in \mathbb{Z}[I]$. Thus the weight of Hilbert modular forms is given by a g-tuple k of integers for $g = [F : \mathbb{Q}]$.

We usually use the symbol PG for G/Z in place of G^{ad} in order to avoid confusion with the adjoint group of the derived semi-simple group G_1 of G,

although in the description of Shimura varieties, we use G^{ad}, which is more common in papers treating Shimura varieties.

Here is some more notation generally used in this chapter. The real Lie group $G(\mathbb{R})$ is isomorphic to $GL_2(\mathbb{R})^I$ by an isomorphism sending $\alpha \in G(\mathbb{Q}) = GL_2(F)$ to a tuple $(\alpha^\sigma)_{\sigma\in I}$ of conjugates. The identity-connected component $G(\mathbb{R})^+$ is made up of g-tuples of matrices $(a_\sigma)_{\sigma\in I}$ with $\det(a_\sigma) > 0$ for all $\sigma \in I$. We put $G(\mathbb{Q})^+ = G(\mathbb{R})^+ \cap G(\mathbb{Q})$. We later define in Section 4.2.2 a slightly different group $G(\mathbb{Q})_+$ for more general reductive groups G, but in this Hilbert modular case, we actually have $G(\mathbb{Q})_+ = G(\mathbb{Q})^+$; so we use the symbol $G(\mathbb{Q})_+$ instead of $G(\mathbb{Q})^+$.

For any lattice $\mathfrak{a} \subset F$, we define $\mathfrak{a}^* = \{x \in F | \mathrm{Tr}_{F/\mathbb{Q}}(\mathfrak{a}x) \subset \mathbb{Z}\}$ and $\mathfrak{d}^{-1} = O^*$ (so, $\mathfrak{d}$ is the absolute different of $F/\mathbb{Q}$). As before, $\mathcal{W}$ is a discrete valuation ring inside $\overline{\mathbb{Q}}$ (on which the embedding $i_p : \overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_p$ is p-adically continuous) and $W = \varprojlim_n W_n$ for $W_n = \mathcal{W}/p^n\mathcal{W}$, which is the p-adic completion of $\mathcal{W}$.

4.1 Hilbert–Blumenthal Moduli

We first describe basic definitions and properties of abelian varieties with real multiplication (AVRM), which are also called by a different name, Hilbert–Blumenthal abelian schemes, and are a direct generalization of elliptic curves, because their Tate modules supply us with p-adic Galois representations into $GL(2)$. Then we describe some features of the moduli of AVRM necessary to prove VCT, putting the construction of the moduli scheme off to Remark 6.21 and Section 7.1.

4.1.1 Abelian Variety with Real Multiplication

Let $A_{/S}$ be an abelian scheme. Here a group scheme over a base scheme S is called an *abelian scheme* if it is a proper smooth geometrically irreducible group scheme over S (the words "geometrically irreducible" mean that its fiber at every geometric point of S is irreducible). A one-dimensional abelian scheme is an elliptic curve. We consider the following Picard functor from $SCH_{/S}$ into $SETS$: $\mathrm{Pic}_{A/S}(S') = \mathrm{Pic}(A_{S'})/\pi_{S'}^*\,\mathrm{Pic}(S') = \mathrm{Ker}(\mathbf{0}_{S'}^*)$ for $\pi : A \to S$. If $f : A \to Y$ is a morphism of abelian schemes, $\mathcal{L} \mapsto f^*\mathcal{L}$ for an invertible sheaf $\mathcal{L}$ induces the dual morphism ${}^t f : \mathrm{Pic}_{Y/S} \to \mathrm{Pic}_{A/S}$. Thus $\mathrm{Pic}_{A/S}$ is a contravariant functor. It is known that $\mathrm{Pic}_{A/S}$ is represented by a (locally Noetherian) reduced group scheme ([FGA] Exp. 232, 1961/62, or [NMD] Chapter 8, or [ABV] Section 13, or [DAV] 1.9). Let ${}^t A$ be the identity-connected component $\mathrm{Pic}^\circ_{A/S}$ of the group scheme representing $\mathrm{Pic}_{A/S}$. Then ${}^t A_{/S}$ is an abelian scheme (see [ABV] Section 13). We write ${}^t f$ again for the dual morphism ${}^t f : {}^t Y \to {}^t A$ of a morphism $f : A \to Y$.

Let $F/\mathbb{Q}$ be a totally real finite extension unramified at the fixed prime p. Write O for the integer ring of F. We consider the following fibered category $\mathcal{A}_F$ of abelian schemes over the category of schemes. Here a *fibered category* $\mathcal{C}$

over a base category $\mathcal{C}_0$ (cf. [SGA] VI) means that we have a specified covariant functor (fiber functor) $\mathcal{C} \to \mathcal{C}_0$. An object of $\mathcal{A}_F$ is a triple $(A, \lambda, \iota)_{/S}$, where

(rm1) $\iota = \iota_A : O \hookrightarrow \mathrm{End}(A_{/S})$ is an embedding of algebras taking the identity to the identity;

(rm2) λ is an O-linear symmetric isogeny $\lambda : A \to {}^tA$ induced by an ample line bundle fiber-by-fiber geometrically (see [GIT] 6.2). Here λ is called *symmetric* if $\lambda = {}^t\lambda$. Such a λ is called a *polarization* of A;

(rm3) The image of ι_A is stable under the Rosati involution on the endomorphism algebra $\mathrm{End}(A) \otimes_{\mathbb{Z}} \mathbb{Q}$: $\alpha \mapsto \alpha^* = \lambda^{-1} \circ {}^t\alpha \circ \lambda$;

(rm4) As $O \otimes_{\mathbb{Z}} \mathcal{O}_S$-modules, we have an isomorphism: $Lie(A) \cong O \otimes_{\mathbb{Z}} \mathcal{O}_S$ ($\Leftrightarrow \pi_*(\Omega_{A/S}) \cong \mathfrak{d}^{-1} \otimes_{\mathbb{Z}} \mathcal{O}_S$ with $\pi : A \to S$ for the absolute different $\mathfrak{d}$ of F) locally under the Zariski topology of S, where the sheaf $Lie(A)$ of Lie algebras of A (i.e., the direct image of the tangent bundle over $A_{/S}$) is an O-module by the action induced from ι.

The fiber functor is given by $\mathcal{A}_F \ni (A, \lambda, \iota)_{/S} \mapsto S \in SCH$.

Here is a more detailed description of the above four conditions. Either by localizing at a rational prime or by extending the scalar to an étale extension, any projective O-module becomes free; so, we often pretend that the isomorphism in (rm4) is global, although $Lie(A)$ is required to be only a locally free module over $(O \otimes_{\mathbb{Z}} \mathcal{O}_S)$ of rank 1. Since an O-module free of finite rank is determined by the characteristic polynomials in $\mathcal{O}_S[t]$ of multiplication by $\alpha \in O$, (rm4) is also equivalent to

(det) *The characteristic polynomial of each $\alpha \in O$ on $Lie(A)$ over $\mathcal{O}_S$ is given by the image of $\prod_\sigma (t - \sigma(\alpha)) \in \mathbb{Z}[t]$ in $\mathcal{O}_S[t]$, where σ runs over all embeddings of F into $\overline{\mathbb{Q}}$.*

A morphism $f : (A, \lambda, \iota)_{/S} \to (A', \lambda', \iota')_{/S}$ of $\mathcal{A}_F$ is an O-linear morphism $f : A_{/S} \to A'_{/S}$ of abelian schemes over S with $\lambda = f^*\lambda' := {}^tf \circ \lambda' \circ f$.

A polarization is an O-linear isogeny $\lambda = \lambda_L : A \to {}^tA$ induced étale locally by a symmetric line bundle $L_{/A}$; that is, we have $(-1)^*L \cong L$ and $\lambda_L(x) = T_x^*(L) \otimes L^{-1}$, where $T_x(y) = x + y$. This definition is equivalent to Mumford's definition ([GIT] 6.3) requiring that λ be induced by a symmetric line bundle L_{s/A_s} for each geometric point $s \in S$ (see [DAV] I.1.6). Indeed, for the universal Poincaré bundle $\mathcal{P}$ on $A \times_S {}^tA_{/A}$, 2λ is associated with $(1 \times \lambda)^*\mathcal{P}$ globally over S (see [GIT] Proposition 6.10).

The morphism λ is an isogeny if and only if L is ample (cf. [ABV] Section 6). We have $\lambda_{L \otimes L'} = \lambda_L + \lambda_{L'}$, and the monoid of $\mathrm{Hom}(A, {}^tA)$ generated by polarizations forms a cone $P(A)$. If $\lambda : A \to {}^tA$ is a polarization, $\mathrm{Ker}(\lambda)$ is given by $A[\mathfrak{c}^{-1}]$ for an integral ideal $\mathfrak{c}^{-1} \neq 0$, because $\mathrm{Ker}(\lambda)$ is self-dual under Cartier duality. Then λ induces ${}^tA \cong A \otimes \mathfrak{c}$ for the following reason. Tensoring A over O with the exact sequence $0 \to O \to \mathfrak{c} \to \mathfrak{c}/O \to 0$, by the divisibility of A, we get another exact sequence $0 \to \mathrm{Tor}_1(\mathfrak{c}/O, A) \to A \to A \otimes_O \mathfrak{c} \to 0$ of *fppf* abelian sheaves (see Section 8.2.1 for *fppf* abelian sheaves). Since O is a Dedekind domain, we have $\mathfrak{c}/O \cong O/\mathfrak{c}^{-1}$, $A[\mathfrak{c}^{-1}] = \mathrm{Tor}_1(\mathfrak{c}/O, A)$

(see [LFE] 1.1.1) and ${}^tA \cong A/A[\mathfrak{c}^{-1}] \cong A \otimes_O \mathfrak{c}$. Such a polarization is often called a $\mathfrak{c}$-*polarization*. Starting from an O-linear polarization λ, we write $\mathfrak{c}(\lambda)$ for its polarization ideal; so, $\ker(\lambda) = A[\mathfrak{c}(\lambda)^{-1}]$ if $\mathfrak{c}(\lambda)^{-1}$ is integral. If S is a $\mathbb{Q}$-scheme, the module $Lie(A)$ is a faithful module over $\mathrm{End}_F^{\mathbb{Q}}(A) = \mathrm{End}_O(A) \otimes_{\mathbb{Z}} \mathbb{Q}$. An element $\phi \in \mathrm{End}_F^{\mathbb{Q}}(A)$ is called symmetric if $\phi^* = \phi$ for the Rosati involution $*$ in (rm3). We write $\mathrm{End}_{F\text{-sym}}^{\mathbb{Q}}(A)$ for the F-subspace of F-linear symmetric endomorphisms in $\mathrm{End}_F^{\mathbb{Q}}(A)$. We put $\mathrm{End}_{O\text{-sym}}(A) = \mathrm{End}_{F\text{-sym}}^{\mathbb{Q}}(A) \cap \mathrm{End}_O(A)$. In particular, the algebra of symmetric endomorphisms $\mathrm{End}_{F\text{-sym}}^{\mathbb{Q}}(A)$ is equal to the image of ι_A and is of dimension 1, which implies $\mathrm{End}_{O\text{-sym}}(A) = O$. Therefore if λ is a $\mathfrak{c}$-polarization, $\mathrm{Hom}_{O\text{-sym}}(A, {}^tA) = \mathrm{End}_{O\text{-sym}}(A) \otimes \mathfrak{c} = \mathfrak{c}$. For $a \in O$, we see easily that $a \circ T_x = T_{a(x)} \circ a$. From this, we have $\lambda_{a^*L} = a^2\lambda_L$. The monoid ring of totally positive elements in the ideal $\mathfrak{c}$ is generated over $\mathbb{N}$ (additively) by square elements. If λ and λ' are two O-linear polarizations, we have $\lambda' = {}^t\iota(a)\lambda$ for $a \in F^\times$. By the positivity of the Rosati involution (i.e., the positive definiteness of the quadratic form $\alpha \mapsto \mathrm{Tr}(\alpha\alpha^*)$ on $\mathrm{End}(A) \otimes \mathbb{Q}$; see [ABV] Section 21), we find that all conjugates of a in $\mathbb{R}$ are positive (i.e., a is *totally positive*: $a \gg 0$). Hence we have a canonical identification $P(A) \cong \mathfrak{c}_+$ for the monoid $\mathfrak{c}_+$ of totally positive elements in $\mathfrak{c}$. Write $F_+^\times \subset F$ for the group of totally positive elements. Then the class of polarizations $\overline{\lambda} = \{\lambda \circ \iota(\alpha) | \alpha \in F_+^\times\}$ is determined by the pair $(\mathfrak{c}, \mathfrak{c}_+ = P(A))$ modulo multiplication by $F_+^\times$ and hence only depends on the strict ideal class of $\mathfrak{c}$ (because $({}^t\iota(\alpha)) \circ \lambda = \lambda \circ \iota(\alpha)$ is an $\alpha^{-1}\mathfrak{c}$-polarization for $\alpha \in F_+^\times$).

We can extend the above definition of the polarization class to a smaller subclass $\overline{\lambda} = \{\lambda \circ \iota(\alpha) | \alpha \in O_{(\Sigma)+}^\times\}$ for a set of rational primes Σ, requiring λ to have degree prime to Σ. Here $\mathbb{Z}_{(\Sigma)} = \mathbb{Q} \cap \mathbb{Z}_\Sigma$ for $\mathbb{Z}_\Sigma = \prod_{p \in \Sigma} \mathbb{Z}_p$, $O_{(\Sigma)} = O \otimes_{\mathbb{Z}} \mathbb{Z}_{(\Sigma)}$, and $O_{(\Sigma)+}^\times = F_+^\times \cap O_{(\Sigma)}$. Shimura was the first to consider such a polarization class, which he called a weak polarization on a given abelian variety (cf. [Sh4] 4.7–8). This is natural since Shimura varieties classify abelian varieties up to isogenies (as we show later in Section 4.2); so, polarization cannot be kept without having some leeway. Suppose that A is defined over a $\mathbb{Z}_{(\Sigma)}$-scheme. Thus $\lambda \in \overline{\lambda}$ is an étale isogeny; so, for each geometric point $s \in S$, the algebraic fundamental group $\pi_1(S, s)$ (in Section 4.4.2) acts on $\overline{\lambda}$. We say that $\overline{\lambda}$ is defined over S if it is stable under the action of $\pi_1(S, s)$ for all geometric points $s \in S$. Changing the geometric point s changes $\pi_1(S, s)$ by an isomorphism (see Section 4.4.2), and we only need to require the stability taking one geometric point on each connected component of S (for more details of this type of results, see Section 6.4.1 where the rationality of level structures is dealt with by such an argument).

Suppose that $\overline{\lambda}$ is defined over S. Then we find an irreducible étale Galois covering S'/S so that we have a member $\lambda : A \times_S S' \to {}^tA \times_S S'$ in $\overline{\lambda}$. The map $\mathrm{Gal}(S'/S) \ni \sigma \mapsto \lambda^{1-\sigma} \in O_{(\Sigma)+}^\times$ is a homomorphism of the finite group $\mathrm{Gal}(S'/S)$ into the torsion-free module $O_{(\Sigma)+}^\times$ (on which $\mathrm{Gal}(S'/S)$

acts trivially). Thus $\lambda^{1-\sigma} = 1$, and λ is defined over S. By this descent argument, we can always find a member $\lambda \in \overline{\lambda}$ that is defined globally over S if $\overline{\lambda}$ is defined over S. Thus our definition of S-integrality is equivalent to having a member λ defined over S in the class $\overline{\lambda}$ (which is the definition of integrality Kottwitz used in [Ko] Section 5). As already remarked, the isogeny 2λ is induced globally by a line bundle; so, if $2 \notin \Sigma$, we have λ in $\overline{\lambda}$ globally associated with an ample line bundle over S.

Take a geometric point $s \in S$ of the base scheme of an AVRM $A_{/S}$. We now study the Tate module $\mathcal{T}(A) = \mathcal{T}_s(A) = \varprojlim_N A[N](k(s))$ for the residue field $k(s)$ of s (which is algebraically closed). By ι, $\mathcal{T}(A)$ is an $\widehat{O}$-module. We put $\mathcal{T}(A)_p = \mathcal{T}(A) \otimes_{\widehat{O}} O_p$ ($O_p = \varprojlim_n O/p^nO$) for a rational prime p. An abelian scheme $A_{/S}$ of relative dimension g is called *ordinary* (at p) if we have an embedding $\mu_p^g \hookrightarrow A[p]$ of finite flat group schemes étale locally.

Proposition 4.1 *Suppose that $k(s)$ is of characteristic p. Let $A_{s/k(s)}$ be the fiber of A at s, and suppose that A_s is* ordinary*; that is, $A_s[p] \cong \mu_p^g \times (\mathbb{Z}/p\mathbb{Z})^g$ for $g = \dim(A_s)$. Then we have $\mathcal{T}(A_s)_p \cong O_p$ as O-modules.*

Proof. Since A_s and the connected component $A_s[p]^\circ \cong \mu_p^g$ have the tangent space of dimension g over $k(s)$, they share the tangent space $Lie(A_s)$ at the origin. As O-modules, they are free of rank 1 over $O \otimes_{\mathbb{Z}} k(s)$ by (rm4). Write $A_s[p]^\circ = Spec(R)$ for a $k(s)$-bialgebra R (e.g., [GME] 1.6.3 for bialgebras). Then for its unique maximal ideal $\mathfrak{m} \subset R$, we have $Lie(A_s) = \mathrm{Hom}_{k(s)}(\mathfrak{m}/\mathfrak{m}^2, k(s))$. By Cartier duality between $A[p]$ and ${}^tA[p]$ (e.g., [GME] 1.7 and Theorem 4.1.17 (2)), we have

$$
\begin{aligned}
&{}^tA_s[p]^{et} \cong \mathrm{Hom}_{\text{Gp-sch}}(A_s[p]^\circ, \mu_p) \hookrightarrow \mathrm{Hom}_{SCH}(A_s[p]^\circ, \mu_p) \cong \\
&\mathrm{Hom}_{k(s)-alg}(k(s)[t]/(t^p), R) \twoheadrightarrow \mathrm{Hom}_{k(s)-alg}(k(s)[t]/(t^p), R/\mathfrak{m}^2) \cong \mathfrak{m}/\mathfrak{m}^2.
\end{aligned}
$$

Since $A[p]^\circ \cong \mu_p^g$ over $k(s)$ for $g = \dim A_s$, it is easy to see that the above morphism induces ${}^tA_s[p]^{et} \otimes_{\mathbb{F}_p} k(s) \cong H^0(A_s, \Omega_{A_s/k(s)})$. By duality and polarization, we get $A_s[p]^{et} \otimes_{\mathbb{F}_p} k(s) \cong Lie(A_s)$. This shows that

$$
Lie(A_s) \cong \mathcal{T}(A_s)_p \otimes_{\mathbb{Z}} k(s) \quad \text{as } O \otimes_{\mathbb{Z}} k(s)\text{-modules.} \tag{4.1}
$$

Then by Nakayama's lemma ([CRT] Theorem 2.2), we conclude from (rm4) the desired assertion. □

The fact (4.1) shows that $Lie(A)_{/S}$ and $\mathcal{O}_S \otimes \mathcal{T}(A[p^\infty]^{et})$ for the locally constant sheaf $\mathcal{T}(A[p^\infty]^{et})_{/S} = \varprojlim_n A[p^n]^{et}$ have the same stalk everywhere if A/S is ordinary at every point over S and p is locally nilpotent over S, and hence they are isomorphic.

Corollary 4.2 *If p is locally nilpotent over S and an abelian scheme $A_{/S}$ is ordinary, we have a canonical isomorphism $Lie(A)_{/S} \cong \mathcal{O}_S \otimes_{\mathbb{Z}} \mathcal{T}(A[p^\infty]^{et})$. If furthermore, $p^n = 0$ over S, we have $Lie(A)_{/S} \cong \mathcal{O}_S \otimes_{\mathbb{Z}} A[p^n]^{et}$, where $A[p^n]^{et}$ is the maximal étale quotient of $A[p^n]$.*

Proposition 4.3 *Suppose that $k(s)$ is of characteristic 0. Then we have an isomorphism $\mathcal{T}(A_s)_p \cong O_p^2$ as O-modules.*

Proof. Since $p^n \in \mathrm{End}(A_s)$ acts on $\Omega_{A_{s/k(s)}}$ by multiplication by p^n, $p^n \in \mathrm{End}(A_s)$ induces an automorphism of the cotangent space; so, $p : A_s \to A_s$ is étale of degree p^{2g}; so, $A_s[p^n] \cong (\mathbb{Z}/p^n\mathbb{Z})^{2g}$, and hence $\mathcal{T}(A_s)_p \cong \mathbb{Z}_p^{2g}$ on which O acts as linear endomorphisms. Thus it is an O_p-module for $O_p = O \otimes_{\mathbb{Z}} \mathbb{Z}_p$. Since the characteristic polynomial of the action of $\alpha \in O$ is given by the square of the characteristic polynomial of the regular representation of α acting on the $\mathbb{Q}$-vector space F (cf. [GME] Theorem 4.1.20), we find that $\mathcal{T}(A_s)_p \cong O_p^2$. □

By the two propositions, we have

Corollary 4.4 *For an AVRM A over S, the kernel $A[N]$ of the multiplication by a positive integer N is a locally free group scheme of rank N^{2g}. If N is invertible on S, $A[N] \cong (O/NO)^2$ étale locally. If the reduced part S^{red} is of characteristic p and A is ordinary over S, we have $A[p^n]^\circ \cong \mathfrak{d}^{-1} \otimes_{\mathbb{Z}} \mu_{p^n}$ and $A[p^n]^{et} \cong O/p^nO$ étale locally. The exact sequence $0 \to A[p^n]^\circ \to A[p^n] \to A[p^n]^{et} \to 0$ is split over each geometric point of S.*

Proof. The multiplication by $N : A \to A$ is a finite flat morphism, because it is quasi-finite and $A_{/S}$ is smooth proper and surjective (cf. [GME] 2.8.1). By a result of Mumford, we show in Sections 4.1.2 and 4.1.6 that the coarse moduli scheme of an AVRM contains a characteristic 0 geometric point and an ordinary characteristic p point. The above two propositions then show that $A[\ell^n]$ for all primes ℓ is locally free of rank ℓ^{2gn} over S. Since $A[N] = \prod_\ell A[\ell^{e(\ell)}]$ for the prime decomposition $N = \prod_\ell \ell^{e(\ell)}$, $A[N]$ is locally free of rank N^{2g}.

Suppose that S^{red} is of characteristic p. We have an exact sequence of locally free group schemes, $0 \to A[p^n]^\circ \to A[p^n] \to A[p^n]^{et} \to 0$, which is split over any perfect field because of $A[p^n]^{red} \cong A[p^n]^{et}$ (see [ABV] Section 14). In particular, it splits over an algebraically closed field and over geometric points. We know that $A[p^n]^{et}_{/S} \cong (O/p^nO)_{/S}$ étale locally by Proposition 4.1. The Cartier dual of $A[p^n]^\circ$ is ${}^tA[p^n]^{et} \cong (O/p^nO)_{/S}$ (e.g., [GME] Theorem 4.1.17), and the dual of $O/NO = (\mathbb{Z}/N\mathbb{Z}) \otimes_{\mathbb{Z}} O$ is given by $\mathfrak{d}^{-1} \otimes_{\mathbb{Z}} \mu_N$ by the pairing induced by the canonical isomorphism $\mathrm{Hom}_{\text{Gp-sch}}(\mu_N, \mathbb{G}_m) = \mathbb{Z}/N\mathbb{Z}$ combined with the trace pairing $\mathfrak{d}^{-1} \times O \to \mathbb{Z}$. □

4.1.2 Moduli Problems with Level Structure

We consider the fibered category of quadruples

$$(A, \lambda, \iota, \phi = \phi_N : (\mathfrak{d}^{-1} \otimes_{\mathbb{Z}} \mu_N) \oplus (O \otimes_{\mathbb{Z}} N^{-1}\mathbb{Z}/\mathbb{Z}) \cong A[N])$$

over a $\mathbb{Z}[\frac{1}{N}]$–SCH. Here the triple $(A, \lambda, \iota : O \hookrightarrow \mathrm{End}(A))$ is an object of $\mathcal{A}_F$, the *level $\Gamma(N)$-structure* ϕ_N is supposed to be O-linear, and the duality

pairing $e_N : A[N] \times {}^tA[N] \to \mu_N$ (the so-called Weil pairing; see (P1) in 8.2.3) composed with the polarization $\lambda : A \to {}^tA$ agrees via ϕ with the natural pairing $\langle \cdot, \cdot \rangle_N : (\mathfrak{d}^{-1} \otimes_{\mathbb{Z}} \mu_N) \oplus (O \otimes_{\mathbb{Z}} N^{-1}\mathbb{Z}/\mathbb{Z}) \to \mu_N$ given by $(a \otimes \zeta, b \otimes m) \mapsto \mathbf{e}\left(\mathrm{Tr}_{F/\mathbb{Q}}(ab)\right)\zeta^m$, where $\mathbf{e}(x) = \exp(2\pi i x)$. We often simply write (A, λ) for $(A, \lambda, \iota : O \hookrightarrow \mathrm{End}(A))$, dropping ι as we agree to have ι always in this chapter. A morphism $f : (A, \lambda, \phi)_{/S} \to (A', \lambda', \phi')_{/S}$ is induced by a morphism $f : (A, \lambda) \to (A', \lambda')$ in $\mathcal{A}_F$ with the additional requirement that $\phi' = f \circ \phi$. We fix a polarization ideal $\mathfrak{c}$, and assume that λ is a $\mathfrak{c}$-polarization. Then for a given $\mathfrak{c}$, the Hilbert–Blumenthal scheme $\mathfrak{M}(\mathfrak{c}, \Gamma(N))_{/\mathbb{Z}[\frac{1}{N}]}$ represents the following functor $\mathcal{E}_{\Gamma(N)} = \mathcal{E}_{\mathfrak{c},\Gamma(N)} : \mathbb{Z}[\frac{1}{N}]\text{-}SCH \to SETS$ given by

$$\mathcal{E}_{\Gamma(N)}(S) = \left[(A, \lambda, \phi_N)_{/S} \ \text{ with } \mathfrak{c}\text{-polarization } \lambda\right], \qquad (\Gamma(N))$$

where as before $[*]$ is the set of isomorphism classes of the objects $*$ inside the brackets. When $N \geq 3$, as we show in Remark 6.21, this functor is representable by a quasi-projective scheme over $\mathbb{Z}[\frac{1}{N}]$ (which is smooth over $\mathbb{Z}[C, \frac{1}{Nd(F)C}]$ for $0 < C \in \mathbb{Q}$ with $(C) = \mathbb{Q} \cap \mathfrak{c}$, as we show later in Section 7.1), and for small $N \leq 2$, we have at least a *coarse moduli* scheme (cf. [GME] 2.3.2).

Let $\mathfrak{N}$ be a nonzero ideal of O. Abusing the notation, we write $\mu_{\mathfrak{N}/\mathbb{Z}}$ for the locally free group scheme (of finite rank) given by the $\mathfrak{N}$-torsion points of $\mathbb{G}_m \otimes_{\mathbb{Z}} \mathfrak{d}^{-1}$: $\mu_{\mathfrak{N}}(R) = \left\{x \in \mathbb{G}_m(R) \otimes_{\mathbb{Z}} \mathfrak{d}^{-1} \middle| \mathfrak{N}x = 0\right\}$. If $\mathfrak{N}$ is generated by a positive integer N, we have $\mu_{(N)} = \mu_N \otimes_{\mathbb{Z}} \mathfrak{d}^{-1}$. We consider the moduli problem classifying test objects $(A, \lambda, i : \mu_{\mathfrak{N}} \hookrightarrow A[N])_{/S}$ of level $\Gamma_1^1(\mathfrak{N})$ in place of the level $\Gamma(N)$-structure. The classifying functor $\mathcal{E}_{\Gamma_1^1(\mathfrak{N})} = \mathcal{E}_{\mathfrak{c},\Gamma_1^1(\mathfrak{N})} : \mathbb{Z}[\frac{1}{N}]\text{-}SCH \to SETS$ for an integer N with $\mathfrak{N} \cap \mathbb{Z} = (N)$ is given by

$$\mathcal{E}_{\Gamma_1^1(\mathfrak{N})}(S) = \left[(A, \lambda, i_{\mathfrak{N}})_{/S} \ \text{ with } \mathfrak{c}(\lambda) = \mathfrak{c}\right], \qquad (\Gamma_1^1(\mathfrak{N}))$$

where $i_{\mathfrak{N}} : \mu_{\mathfrak{N}} \hookrightarrow A[N]$ is the inclusion of locally free group schemes. If each test object $(A, \lambda, i_{\mathfrak{N}})_{/S}$ is rigid without nontrivial automorphism, the functor is representable by a quasi-projective scheme $\mathfrak{M}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))_{/\mathbb{Z}[\frac{1}{N}]}$, which is smooth over $\mathbb{Z}[\frac{1}{Nd(F)}]$. Thus the functor $\mathcal{E}_{\Gamma_1^1(\mathfrak{N})}$ is representable if $\mathfrak{N}$ is deep enough. The coarse moduli $\mathfrak{M}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))_{/\mathbb{Z}[\frac{1}{N}]}$ exist for all $\mathfrak{N}$.

Let $O_+^\times = F_+^\times \cap O^\times$. We may ease the polarization condition, that is, replace a polarization λ by a polarization class $O_+^\times \lambda = \{\epsilon\lambda | \epsilon \in O_+^\times\}$. Easing the polarization requirement allows us to change morphisms of our category from O-linear isomorphisms to O-linear isomorphisms up to unit factors (we later study categories up to a larger class of isogenies in Section 4.2). For a given $O_+^\times \lambda$ of A, we say that $O_+^\times \lambda$ is defined over S if we find an ample line bundle, locally under Zariski topology on S, giving rise to an element in $O_+^\times \lambda$ (cf. [ABV] Section 8). Consider the following functor $\mathcal{E}_{\Gamma_1(\mathfrak{N})} : SCH_{/\mathbb{Z}[1/N]} \to SETS$ (of level $\Gamma_1(\mathfrak{N})$) for an integer N with $\mathfrak{N} \cap \mathbb{Z} = (N)$ given by

$$\mathcal{E}_{\Gamma_1(\mathfrak{N})}(S) = \left[(A, O_+^\times \lambda, i_{\mathfrak{N}})_{/S} \ \text{ with } \mathfrak{c}\text{-polarization class } O_+^\times \lambda\right], \qquad (\Gamma_1(\mathfrak{N}))$$

where $(A, O_+^\times\lambda, i) \cong (A', O_+^\times\lambda', i')$ if we have an isomorphism $f : A \cong A'$ of abelian schemes with $f \circ i = i'$ and $O_+^\times\lambda = O_+^\times f^*\lambda'$.

Since $\epsilon \in O_+^\times = O^\times \cap F_+^\times$ with $\epsilon \equiv 1 \mod \mathfrak{N}$ gives an automorphism of $(A, O_+^\times\lambda, i_\mathfrak{N})_{/S}$, the automorphism group $\mathrm{Aut}((A, O_+^\times\lambda, i_\mathfrak{N})_{/S})$ is not trivial; so, the functor $\mathcal{E}_{\Gamma_1(\mathfrak{N})}$ is not representable in the strict sense that we have a moduli scheme $\mathfrak{M}(\mathfrak{c}, \mathfrak{N})$ and a universal couple $(\mathbb{X}, O_+^\times\boldsymbol{\lambda}, \mathbf{i}_\mathfrak{N})$ such that for any triple $(A, O_+^\times\lambda, i_\mathfrak{N})_{/S} \in \mathcal{E}_{\Gamma_1(\mathfrak{N})}(S)$, we have a unique morphism $f : S \to \mathfrak{M}(\mathfrak{c}, \mathfrak{N})$ such that $f^*(\mathbb{X}, O_+^\times\boldsymbol{\lambda}, \mathbf{i}_\mathfrak{N})$ is isomorphic in a unique way to $(A, O_+^\times\lambda, i_\mathfrak{N})_{/S}$. But if $\mathfrak{N}$ is sufficiently deep, the uniqueness is valid up to a multiple of $O_+^\times$. The coarse moduli scheme $\mathfrak{M}(\mathfrak{c}, \mathfrak{N})$ with functorial isomorphisms $\mathfrak{M}(\mathfrak{c}, \mathfrak{N})(k) \cong \mathcal{E}_{\Gamma_1(\mathfrak{N})}(k)$ (for algebraically closed fields k) exists always. This fact can be seen as follows. Letting $\epsilon \in O_+^\times$ act on $\mathcal{E}_{\mathfrak{c},\Gamma_1^1(\mathfrak{N})}$ by $(A, \lambda, i) \mapsto (A, \epsilon\lambda, i)$, we find that $\mathcal{E}_{\Gamma_1(\mathfrak{N})}$ has its coarse moduli if $\mathcal{E}_{\Gamma_1^1(\mathfrak{N})}$ has, and the coarse moduli scheme $\mathfrak{M}(\mathfrak{c}, \mathfrak{N})_{/\mathbb{Z}[\frac{1}{N}]}$ is given by $\mathfrak{M}(\mathfrak{c}, \mathfrak{N}) = O_+^\times\backslash\mathfrak{M}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))$. As is clear from this construction, for each member $(A, O_+^\times\lambda, i_\mathfrak{N})_{/S} \in \mathcal{E}_{\Gamma_1(\mathfrak{N})}(S)$, we have a unique morphism $f : S \to \mathfrak{M}(\mathfrak{c}, \mathfrak{N})$ with $f^*(\mathbb{X}, O_+^\times\boldsymbol{\lambda}, \mathbf{i}_\mathfrak{N}) \cong (A, O_+^\times\lambda, i_\mathfrak{N})_{/S}$, but this last isomorphism is unique modulo $O_+^\times$ even if $(\mathbb{X}, O_+^\times\boldsymbol{\lambda}, \mathbf{i}_\mathfrak{N})_{/\mathfrak{M}(\mathfrak{c},\mathfrak{N})}$ exists (e.g., if $\epsilon = \varepsilon^2$ with $\varepsilon \in O^\times$ as long as $\epsilon \equiv 1 \mod \mathfrak{N}$ for $\epsilon \in O_+^\times$). This is a typical phenomenon we encounter if we consider the moduli problem with ambiguity of polarization (which is a moduli problem of $GL(2)$-type). To justify this point, we need to change the structure of the category $\mathcal{A}_F$ declaring some class of isogenies as isomorphisms. We study this point later in Section 4.2. We record what we have described.

Theorem 4.5 *Let the notation and the assumption be as above. Let Γ be one of the level structures of type $\Gamma(N)$, $\Gamma_1^1(\mathfrak{N})$, and $\Gamma_1(\mathfrak{N})$. Then we have*

(1) *The coarse moduli scheme $\mathfrak{M}(\mathfrak{c}, \Gamma)$ representing the functor $\mathcal{E}_\Gamma$ always exists over $\mathbb{Z}\left[\frac{1}{N}\right]$ for the minimal positive integer $N \in \mathfrak{N}$.*
(2) *Take $\mathfrak{c}^{-1} \subset O$, and let C be the minimal positive integer in $\mathfrak{c}^{-1}$. The scheme $\mathfrak{M}(\mathfrak{c}, \Gamma)$ for $\Gamma = \Gamma(N)$ and $\Gamma_1^1(\mathfrak{N})$ is smooth over $\mathbb{Z}[\frac{1}{d(F)NC}]$ if $\mathfrak{N}$ is sufficiently deep that test objects of type Γ have no nontrivial automorphisms (e.g., if $N \geq 3$ for $\Gamma = \Gamma(N)$).*
(3) *The scheme $\mathfrak{M}(\mathfrak{c}, \Gamma)$ for $\Gamma = \Gamma(N)$ and $\Gamma_1^1(\mathfrak{N})$ represents the functor $\mathcal{E}_\Gamma$ over $\mathbb{Z}\left[\frac{1}{N}\right]$ if $\mathfrak{N}$ is sufficiently deep that test objects of type Γ have no nontrivial automorphisms.*

4.1.3 Complex Analytic Hilbert Modular Forms

We describe the Archimedean theory of Hilbert modular forms, limiting ourselves to what we need later. We accommodate the $\Gamma_1(\mathfrak{N})$-type moduli problem for an integral ideal $\mathfrak{N}$ (as well as that of $\Gamma(N)$-type for an integer N). When we deal with $\Gamma(\mathfrak{N})$-structure, the ideal $\mathfrak{N}$ is assumed to be generated

by the integer N. We use the symbol ϕ (resp. i) to indicate a level $\Gamma(N)$-structure (resp. a $\Gamma_1(\mathfrak{N})$-structure). Over $\mathbb{C}$, by the theory of theta functions due to Riemann–Poincaré–Lefschetz (see [ABV] Chapter I), the category of Hilbert modular test objects $(A, \lambda, \phi, (\text{resp. } i), \omega)$ is equivalent to the category of triples $(\mathcal{L}, \lambda, \phi, (\text{resp. } i))$ made of the following data: $\mathcal{L}$ is an O-lattice in $O \otimes_{\mathbb{Z}} \mathbb{C} = \mathbb{C}^I$, an alternating pairing $\lambda : \mathcal{L} \wedge_O \mathcal{L} \cong \mathfrak{c}^*$, and a level N-structure ϕ or i. A level $\Gamma_1(\mathfrak{N})$-structure i is an inclusion $i : \mathfrak{N}^*/O^* \hookrightarrow F\mathcal{L}/\mathcal{L}$, and a level $\Gamma(N)$-structure ϕ is an inclusion $\phi : (\mathfrak{N}^*/O^* \times O/\mathfrak{N}) \hookrightarrow F\mathcal{L}/\mathcal{L}$, where $\mathfrak{N}^* = \mathfrak{N}^{-1}\mathfrak{d}^{-1}$ for the absolute different $\mathfrak{d}$ of F.

The alternating form λ is supposed to be positive in the sense that the symmetric bilinear form $(u, v) \mapsto \frac{\lambda(u,v)}{\mathrm{Im}(uv^c)}$ is totally positive definite and that the pairing $\langle \cdot, \cdot \rangle_N$ matches the pairing

$$N^{-1}\mathcal{L}/\mathcal{L} \bigwedge N^{-1}\mathcal{L}/\mathcal{L} \xrightarrow{\lambda} N^{-1}\mathfrak{c}^*/\mathfrak{c}^* \xrightarrow{\mathbf{e}\circ \mathrm{Tr}_{F/\mathbb{Q}}} \mu_N$$

via ϕ when we deal with level $\Gamma(N)$-structure. The differential ω can be recovered by $\iota : A(\mathbb{C}) = \mathbb{C}^I/\mathcal{L}$ so that $\omega = \iota^* du$, where $u = (u_\sigma)_{\sigma \in I}$ is the variable on $\mathbb{C}^I$. Conversely, for a given test object (A, λ) in $\mathcal{A}_F$ with $\mathfrak{c}$-polarization λ,

$$\mathcal{L}_A = \left\{ \int_\gamma \omega \in O \otimes_{\mathbb{Z}} \mathbb{C} \middle| \gamma \in H_1(A(\mathbb{C}), \mathbb{Z}) \right\}$$

is a lattice in $\mathbb{C}^I$, and the polarization $\lambda : {}^tA \cong A \otimes \mathfrak{c}$ induces $\mathcal{L} \wedge \mathcal{L} \cong \mathfrak{c}^*$ (see (P1-3) in Section 8.2.3).

Using this equivalence, we can relate our geometric definition of Hilbert modular forms with the classical analytic definition. Define $\mathfrak{Z}$ by the product of I copies of the upper half complex plane $\mathfrak{H}$. By $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \mapsto \left(\frac{a^\sigma \sqrt{-1} + b^\sigma}{c^\sigma \sqrt{-1} + d^\sigma}\right)_\sigma$, we have $G_1(\mathbb{R})/C_{\mathbf{i}} \cong \mathfrak{Z}$ for $C_{\mathbf{i}} = SO_2(F \otimes_{\mathbb{Q}} \mathbb{R})$ and $\mathbf{i} = (\sqrt{-1}, \ldots, \sqrt{-1}) \in \mathfrak{Z}$ (which is the origin $\mathbf{0}$ of $\mathfrak{Z}$). We regard $\mathfrak{Z} \subset F \otimes_{\mathbb{Q}} \mathbb{C} = \mathbb{C}^I$ made up of $z = (z_\sigma)_{\sigma \in I}$ with a totally positive imaginary part.

We identify μ_N with $N^{-1}\mathbb{Z}/\mathbb{Z}$ by $\exp(2\pi i m) \leftrightarrow (m \mod \mathbb{Z})$. This identification induces $\mu_{(N)} = \mu_N \otimes \mathfrak{d}^{-1} \cong (N^{-1}\mathbb{Z}/\mathbb{Z}) \otimes \mathfrak{d}^{-1} = (N)^*/O^*$, which further induces $\mu_{\mathfrak{N}} \cong \mathfrak{N}^*/O^*$. Choose two ideals $\mathfrak{a}$ and $\mathfrak{b}$ prime to $\mathfrak{N}$ with $\mathfrak{a}\mathfrak{b}^{-1} = \mathfrak{c}$. For each $z \in \mathfrak{Z}$, we define

$$\mathcal{L}_z = 2\pi\sqrt{-1}(\mathfrak{b}z - \mathfrak{a}^*), \ \lambda_z(2\pi\sqrt{-1}(az - b), 2\pi\sqrt{-1}(cz - d)) = (ad - bc) \in \mathfrak{c}^*$$

with $i_z : \mu_{\mathfrak{N}} = \mathfrak{N}^*/O^* \to \mathbb{C}^I/\mathcal{L}_z$ given by $i_z(a \mod O^*) = (-2\pi\sqrt{-1}a + \mathcal{L}_z)$ and $\phi_{\mathfrak{N},z} : \mathfrak{N}^*/O^* \times \mathfrak{N}^{-1}/O \to \mathbb{C}^I/\mathcal{L}_z$ given by, for $J_1 = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$,

$$\phi_{\mathfrak{N},z}(a, b) \equiv 2\pi\sqrt{-1}(-a + bz) \equiv (a, b)J_1 \cdot {}^t(z, 1) \mod \mathcal{L}_z.$$

Taking association $(a, b) \mapsto -a + bz$ instead of $(a, b) \mapsto az + b$ looks strange (in order to put a complex structure on $L_{\mathbb{R}} = L \otimes_{\mathbb{Z}} \mathbb{R}$ for $L = \mathfrak{a}^* \oplus \mathfrak{b}$). However, we later interpret our moduli scheme in Deligne's language in Section 4.2, and

then $\mathfrak{Z}$ is identified with the set X^+ of all conjugates under $G(\mathbb{R})^+$ of the group homomorphism $h_0 : \mathbb{C}^\times \ni a+bi \mapsto \left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right) \in G(\mathbb{R})^+$. The identification is given by $\mathfrak{Z} \ni g(\mathbf{i}) \mapsto g \cdot h_0 g^{-1} \in X^+$ ($\mathbf{i} = (\sqrt{-1}, \dots, \sqrt{-1}) \in \mathfrak{Z}$). For each $h \in X^+$ (corresponding to $z \in \mathfrak{Z}$), $L_{\mathbb{R}} \ni v \mapsto h(a+bi)v$ gives the complex structure we want. For the alternating pairing $\langle x, y\rangle = {}^t x \cdot J_1 y$ of $V = F^2$ (column vectors), $H_h(x,y) = \langle x, h(i)y\rangle$ gives a positive definite Hermitian form on $L_{\mathbb{R}} = V_{\mathbb{R}}$ so that $\langle \cdot, \cdot \rangle$ gives rise to the Riemann form on $L_{\mathbb{R}}/L = \mathbb{C}^I/\mathcal{L}_z$. The natural action $v \mapsto \alpha v$ of $\alpha \in G(\mathbb{R})^+$ on $L_{\mathbb{R}}$ induces the right action: $z \mapsto \alpha^{-1}(z)$ on $\mathfrak{Z}$. Indeed, we have

$$\langle \alpha x, h(i)\alpha y\rangle = \langle x, \alpha^\iota h(i)\alpha y\rangle \equiv \langle x, \alpha^{-1}h(i)\alpha y\rangle \mod T(\mathbb{R})^+$$

for the adjoint involution ι, since $\alpha^\iota = J_1^{-1} \cdot {}^t\alpha J_1 = \det(\alpha)\alpha^{-1}$. In other words, the natural left action of G on V induces the right action $z \mapsto \alpha^\iota(z) = \alpha^{-1}(z)$ on $\mathfrak{Z}$. This is the explanation of why we need to have $(a,b) \mapsto -a+bz$.

If $\alpha = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in G(\mathbb{Q})_+$, the multiplication by $(-cz+a) \in F_{\mathbb{C}}^\times$ induces an isogeny

$$(\mathbb{C}^I/\mathcal{L}_{\alpha^{-1}(z)}, \lambda, \phi_{\mathfrak{N},\alpha^{-1}(z)}) \xrightarrow{x \mapsto (-cz+a)x} (\mathbb{C}^I/\mathcal{L}_z, \det(\alpha)\lambda, \phi_{\mathfrak{N},z} \circ \alpha). \tag{4.2}$$

This shows that the action of $\alpha \in G(\mathbb{Q})_+$ sends the complex point of the projective limit $\mathfrak{M}(\mathbb{C}) = \varprojlim_N \mathfrak{M}(\mathfrak{c}, \Gamma(N))(\mathbb{C})$ represented by $z \in \mathfrak{Z}$ to that of $\alpha^{-1}(z)$. The above isogeny is an isomorphism if and only if $\alpha \in \Gamma(\mathfrak{N}; \mathfrak{a}, \mathfrak{b})$ for the arithmetic subgroup $\Gamma(\mathfrak{N}; \mathfrak{a}, \mathfrak{b})$ of $G(\mathbb{Q})$ defined below.

For four fractional ideals $\mathfrak{a}_{ij} \subset F$, we write symbolically $\left(\begin{smallmatrix} \mathfrak{a}_{11} & \mathfrak{a}_{12} \\ \mathfrak{a}_{21} & \mathfrak{a}_{22} \end{smallmatrix}\right)$ for the O-lattice in $M_2(F)$ made up of matrices whose (i,j)-entry is in $\mathfrak{a}_{ij}$ for $i,j = 1,2$. Define three types of congruence subgroups of $G(\mathbb{Q})_+$ by

$$\begin{aligned}
\Gamma_1(\mathfrak{N}; \mathfrak{a}, \mathfrak{b}) &= \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \left(\begin{smallmatrix} O & (\mathfrak{ab})^* \\ \mathfrak{Nabd} & O \end{smallmatrix}\right) \middle| ad - bc \in O_+^\times, \quad a-1 \in \mathfrak{N} \right\}, \\
\Gamma_1^1(\mathfrak{N}; \mathfrak{a}, \mathfrak{b}) &= \Gamma_1(\mathfrak{N}; \mathfrak{a}, \mathfrak{b}) \cap SL_2(F), \\
\Gamma(\mathfrak{N}; \mathfrak{a}, \mathfrak{b}) &= \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_1^1(\mathfrak{N}; \mathfrak{a}, \mathfrak{b}) \middle| b \in \mathfrak{N}(\mathfrak{ab})^* \right\}.
\end{aligned} \tag{4.3}$$

Since $\det(\Gamma_1^1(\mathfrak{N}; \mathfrak{a}, \mathfrak{b})) = 1$, the condition $a-1 \in \mathfrak{N}$ for $\Gamma_1^1(\mathfrak{N}; \mathfrak{a}, \mathfrak{b})$ automatically implies $d-1 \in \mathfrak{N}$.

For the moment, we deal with $\Gamma_1(\mathfrak{N})$-structures. We let $g = (g_\sigma) \in G(\mathbb{R})^+$ act on $\mathfrak{Z}$ by linear fractional transformation of g_σ on each component z_σ. It is easy to verify by (4.2), similarly to Theorem 2.39, that

(Ism) $\qquad (\mathcal{L}_z, \lambda_z, i_z) \cong (\mathcal{L}_w, \lambda_w, i_w) \iff w = \gamma(z)$ for $\gamma \in \Gamma_1^1(\mathfrak{N}; \mathfrak{a}, \mathfrak{b})$.

Thus the complex points of the coarse moduli scheme $\mathfrak{M}(\mathfrak{c}; \Gamma_1^1(\mathfrak{N}))$ are given by $\Gamma_1^1(\mathfrak{N}; \mathfrak{a}, \mathfrak{b})\backslash\mathfrak{Z}$. However, it looks odd that we have an extra dependence on the pair $(\mathfrak{a}, \mathfrak{b})$ not just on $\mathfrak{c}$ and $\mathfrak{N}$. Actually, as long as $\mathfrak{ab}^{-1} = \mathfrak{c}$, the complex manifolds $\Gamma_1^1(\mathfrak{N}; \mathfrak{a}, \mathfrak{b})\backslash\mathfrak{Z}$ are isomorphic for all choices of $(\mathfrak{a}, \mathfrak{b})$. To see this fact, let us analyze the equivalence classes of cusps of $\Gamma(O; O, \mathfrak{c}^{-1})$. Cusps of

$\mathfrak{Z} = \mathfrak{H}^I$ are given by $\mathbf{P}^1(F) = F \cup \{\infty\}$. Recall that $G(\mathbb{Q})_+$ is the subgroup of $G(\mathbb{Q})$ made up of matrices with totally positive determinants.

The groups $G_1(\mathbb{Q})$ and $G(\mathbb{Q})_+$ act on $\mathbf{P}^1(F)$ by linear fractional transformation (or equivalently by matrix multiplication on homogeneous coordinates). The stabilizer of the infinity cusp is the subgroup $B \subset G$ of upper triangular matrices. Thus $\mathbf{P}^1(F) \cong G_1(\mathbb{Q})/B_1(\mathbb{Q}) = G(\mathbb{Q})_+/B(\mathbb{Q})_+$ for $B(\mathbb{Q})_+ = B(\mathbb{Q}) \cap G(\mathbb{Q})_+$ and $B_1 = B \cap G_1$. For an arithmetic subgroup $\Gamma \subset G_1(\mathbb{Q})$, its closure $\widehat{\Gamma}$ in $G_1(\mathbb{A}^{(\infty)})$ is an open compact subgroup, and $\Gamma = \widehat{\Gamma} \cap G_1(\mathbb{Q})$. Two cusps s and s' are equivalent under Γ if $\gamma(s) = s'$ for some $\gamma \in \Gamma$. Thus, by the strong approximation theorem, $G_1(\mathbb{Q})$ is dense in $G_1(\mathbb{A}^{(\infty)})$, and the equivalent classes of cusps are in bijection with $\Gamma\backslash G_1(\mathbb{Q})/B_1(\mathbb{Q}) = \widehat{\Gamma}\backslash G_1(\mathbb{A}^{(\infty)})/B_1(\mathbb{Q})$. We now make this set explicit for $\Gamma = \Gamma(O; O, \mathfrak{c}^{-1})$. For $g \in G_1(\mathbb{Q})$, we write $\ell(g)$ for the left column vector of g. Then writing $\ell(g) = {}^t(a, c)$, define $il_{\mathfrak{c}}(g) = (c\mathfrak{c}\mathfrak{d} + aO)$ as a fractional ideal of F. Then $\ell(g)\widetilde{b} = \ell(gb)$ for $b = \left(\begin{smallmatrix} \widetilde{b} & * \\ 0 & * \end{smallmatrix}\right) \in B_1(\mathbb{Q})$. If we choose another base, $il_{\mathfrak{c}}(g) = c'\mathfrak{c}\mathfrak{d} + a'O$, we find $\gamma \in \Gamma(O; O, \mathfrak{c}^{-1})$ with $\gamma \cdot \ell(g) = {}^t(a', c') = \ell(\gamma g)$. Thus by $g \mapsto il_{\mathfrak{c}}(g)$, we find an isomorphism $\Gamma(O; O, \mathfrak{c}^{-1})\backslash G_1(\mathbb{Q})/B_1(\mathbb{Q}) \cong Cl_F$ for the class group Cl_F of F. In other words, the cusps of $\Gamma(O; O, \mathfrak{c}^{-1})$ are in bijection with the class group of F. We can extend $il_{\mathfrak{c}}$ to $G_1(\mathbb{A}^{(\infty)})$ by $il_{\mathfrak{c}}(g) = (c\widehat{\mathfrak{c}\mathfrak{d}} + a\widehat{O}) \cap F$, which gives rise to the identity $\widehat{\Gamma}(O; O, \mathfrak{c}^{-1})\backslash G_1(\mathbb{A}^{(\infty)})/B_1(\mathbb{Q}) \cong Cl_F$. In particular, $il_{\mathfrak{c}}(g)\mathfrak{a} = il_{\mathfrak{c}}(ga)$ for $a \in B_1(\mathbb{A}^{(\infty)})$ with $\widetilde{a}\widehat{O} \cap F = \mathfrak{a}$. By the Iwasawa decomposition (e.g., [EPE] Lemma 9.2), we have $G(\mathbb{A}^{(\infty)}) = \widehat{\Gamma}(O; O, \mathfrak{c}^{-1})B_1(\mathbb{A}^{(\infty)})$; so, each cusp of $\Gamma(O; O, \mathfrak{c}^{-1})$ is represented by $b \in B_1(\mathbb{A}^{(\infty)})$. Thus we have an association of cusps s to ideals of F given by choosing $b \in B_1(\mathbb{A}^{(\infty)})$ representing s and sending s to the ideal class of $il_{\mathfrak{c}}(b)$.

There is another description of cusps of $\Gamma(O; O, \mathfrak{c}^{-1})$. The set of pairs $(\mathfrak{a}, \mathfrak{b})$ with $\mathfrak{a}\mathfrak{b}^{-1} = \mathfrak{c}$ is in bijection with the set of cusps of $\Gamma(O; O, \mathfrak{c}^{-1})$ in the following way. Since $\mathfrak{a}\mathfrak{b}^{-1}$ is fixed to be $\mathfrak{c}$, we find $(\mathfrak{a}, \mathfrak{b}) = (\mathfrak{a}, \mathfrak{a}\mathfrak{c}^{-1}) = \mathfrak{a}(O, \mathfrak{c}^{-1})$. Thus a standard choice of a cusp is $(O, \mathfrak{c}^{-1})$, which we call the infinity cusp of $\mathfrak{M}(\mathfrak{c}, 1)$, and $\mathfrak{a}$ corresponds to a cusp s if $il_{\mathfrak{c}}(b) = \mathfrak{a}$ for $b \in B_1(\mathbb{A}^{(\infty)})$ representing s.

For one Γ of these level structures $\Gamma_1^1(\mathfrak{N})$, $\Gamma_1(\mathfrak{N})$, and $\Gamma(N)$, the moduli space $\mathfrak{M}(\mathfrak{c}, \Gamma)$ has a finite number of cusps (depending on the choice of a level Γ-structure on the Tate AVRM $\mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q)$ we define in Section 4.1.5) over a given cusp s represented by $(\mathfrak{a}, \mathfrak{b})$. When $\Gamma_1^1(\mathfrak{N}) \subset \Gamma \subset \Gamma_1(\mathfrak{N})$, we have a canonical level structure $i_{can,\mathfrak{N}}$ as we specify in the following subsection. When $\Gamma = \Gamma(N)$, we have a canonical level structure $\phi_{can,N}$ over an extension ring $\mathbb{Z}[\frac{1}{d(F)N}]((N^{-1}\mathfrak{a}\mathfrak{b}))$ of the canonical ring of definition $\mathbb{Z}[\frac{1}{d(F)N}]((\mathfrak{a}\mathfrak{b}))$ of $\mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q)$. We simply write $(\mathfrak{a}, \mathfrak{b})$ for the cusp of $\mathfrak{M}(\mathfrak{c}, \Gamma)$ with this canonical choice of the level structure. Write

$$\Gamma_1^1(\mathfrak{c}, \mathfrak{N}) = \Gamma_1^1(\mathfrak{N}; O, \mathfrak{c}^{-1}),\ \Gamma_1(\mathfrak{c}, \mathfrak{N}) = \Gamma_1(\mathfrak{N}; O, \mathfrak{c}^{-1}),\ \Gamma(\mathfrak{c}, \mathfrak{N}) = \Gamma(\mathfrak{N}; O, \mathfrak{c}^{-1}).$$

For $\gamma \in SL_2(F)$ with $\gamma \begin{pmatrix} \mathfrak{a}^* \\ \mathfrak{b} \end{pmatrix} = \begin{pmatrix} O^* \\ \mathfrak{c}^{-1} \end{pmatrix}$, γ brings the cusp corresponding to the pair $(\mathfrak{a}, \mathfrak{b})$ to the standard cusp $(O, \mathfrak{c}^{-1})$. In particular, we have

$$\gamma \Gamma_1^1(\mathfrak{c}, \mathfrak{N})\gamma^{-1} = \Gamma_1^1(\mathfrak{N}; \mathfrak{a}, \mathfrak{b}),\ \gamma \Gamma_1(\mathfrak{c}, \mathfrak{N})\gamma^{-1} = \Gamma_1(\mathfrak{N}; \mathfrak{a}, \mathfrak{b}), \quad \text{and}$$
$$\gamma \Gamma(\mathfrak{c}, \mathfrak{N})\gamma^{-1} = \Gamma(\mathfrak{N}; \mathfrak{a}, \mathfrak{b}). \quad (4.4)$$

Thus we find $\Gamma_1^1(\mathfrak{c}, \mathfrak{N})\backslash \mathfrak{Z} \xrightarrow[\gamma]{\sim} \Gamma_1^1(\mathfrak{N}; \mathfrak{a}, \mathfrak{b})\backslash \mathfrak{Z}$ for the above choice of γ. For each ideal $\mathfrak{a}$, $(\mathfrak{a}, \mathfrak{a}\mathfrak{c}^{-1})$ gives another cusp. The two cusps $(\mathfrak{a}, \mathfrak{a}\mathfrak{c}^{-1})$ and $(\mathfrak{s}, \mathfrak{s}\mathfrak{c}^{-1})$ are equivalent under $\Gamma_1(\mathfrak{c}, \mathfrak{N})$ if $\mathfrak{a} = \alpha\mathfrak{s}$ for an element $\alpha \in F^\times$ with $\alpha \equiv 1$ mod $\mathfrak{N}$ in $F_{\mathfrak{N}}^\times$, where $F_{\mathfrak{N}} = \prod_{\mathfrak{l}|\mathfrak{N}} F_{\mathfrak{l}}$ for prime factors $\mathfrak{l}$ of $\mathfrak{N}$.

Returning to the coarse moduli $\mathfrak{M}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))$ over $\mathbb{Z}[\frac{1}{d(F)N}]$ $(\mathfrak{N} \cap \mathbb{Z} = (N))$ for the functor $\mathcal{E}_{\mathfrak{c}, \Gamma_1^1(\mathfrak{N})} : \mathbb{Z}[\frac{1}{d(F)N}]\text{-}SCH \to SETS$,

$$\mathcal{E}_{\mathfrak{c}, \Gamma_1^1(\mathfrak{N})}(S) = [(A, \lambda, i)_{/S} \ \text{ with } \mathfrak{c}(\lambda) = \mathfrak{c}],$$

we have $\mathfrak{M}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))(\mathbb{C}) = \mathcal{E}_{\mathfrak{c}, \Gamma_1^1(\mathfrak{N})}(\mathbb{C}) \cong \Gamma_1^1(\mathfrak{c}, \mathfrak{N})\backslash \mathfrak{Z}$ from (Ism). In exactly the same way, we get $\mathfrak{M}(\mathfrak{c}, \Gamma(N))(\mathbb{C}) \cong \Gamma(\mathfrak{c}, N)\backslash \mathfrak{Z}$ and $\mathfrak{M}(\mathfrak{c}, \mathfrak{N})(\mathbb{C}) \cong \Gamma_1(\mathfrak{c}, \mathfrak{N})\backslash \mathfrak{Z}$, canonically.

Let $G = \mathrm{Res}_{O/\mathbb{Z}} GL(2)$, and set $\widehat{\mathbb{Z}} = \prod_{\ell:\text{prime}} \mathbb{Z}_\ell$. Define four open compact subgroups of $G(\mathbb{A}^{(\infty)})$ by

$$\begin{aligned}
\widehat{\Gamma}_0(\mathfrak{N}) &= \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G(\widehat{\mathbb{Z}}) \middle| c \in \mathfrak{N}\widehat{O} \right\}, \\
\widehat{\Gamma}_1(\mathfrak{N}) &= \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \widehat{\Gamma}_0(\mathfrak{N}) \middle| a \equiv 1 \mod \mathfrak{N}\widehat{O} \right\}, \\
\widehat{\Gamma}(\mathfrak{N}) &= \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \widehat{\Gamma}_1(\mathfrak{N}) \middle| b \in \mathfrak{N}\widehat{O},\ a \equiv 1 \mod \mathfrak{N}\widehat{O} \right\}, \\
\widehat{\Gamma}_1^1(\mathfrak{N}) &= \left\{ x \in \widehat{\Gamma}_1(\mathfrak{N}) \middle| \det(x) \in 1 + \mathfrak{N}\widehat{O} \right\}.
\end{aligned} \quad (4.5)$$

Put

$$S_*(\mathfrak{N}) = \begin{pmatrix} d & 0 \\ 0 & 1 \end{pmatrix}^{-1} \widehat{\Gamma}_*(\mathfrak{N}) \begin{pmatrix} d & 0 \\ 0 & 1 \end{pmatrix} \quad \text{for } * = 0, 1,\ S_1^1(\mathfrak{N}) = \begin{pmatrix} d & 0 \\ 0 & 1 \end{pmatrix}^{-1} \widehat{\Gamma}_1^1(\mathfrak{N}) \begin{pmatrix} d & 0 \\ 0 & 1 \end{pmatrix}$$
$$\text{and } S(\mathfrak{N}) = \begin{pmatrix} d & 0 \\ 0 & 1 \end{pmatrix}^{-1} \widehat{\Gamma}(\mathfrak{N}) \begin{pmatrix} d & 0 \\ 0 & 1 \end{pmatrix} \quad (4.6)$$

for an idele d with $dO = \mathfrak{d}$ and $d^{(\mathfrak{d})} = 1$. Then taking an idele c with $c\widehat{O} = \widehat{\mathfrak{c}}$ and $c^{(\mathfrak{c})} = 1$, we see that

$$\Gamma_1(\mathfrak{c}, \mathfrak{N}) = \begin{pmatrix} c & 0 \\ 0 & 1 \end{pmatrix} S_1(\mathfrak{N}) \begin{pmatrix} c & 0 \\ 0 & 1 \end{pmatrix}^{-1} \cap G(\mathbb{Q})_+,$$
$$\Gamma(\mathfrak{c}, \mathfrak{N}) = \begin{pmatrix} c & 0 \\ 0 & 1 \end{pmatrix} S(\mathfrak{N}) \begin{pmatrix} c & 0 \\ 0 & 1 \end{pmatrix}^{-1} \cap G_1(\mathbb{Q}) \ \text{ and}$$
$$\Gamma_1^1(\mathfrak{c}, \mathfrak{N}) = \begin{pmatrix} c & 0 \\ 0 & 1 \end{pmatrix} S_1^1(\mathfrak{N}) \begin{pmatrix} c & 0 \\ 0 & 1 \end{pmatrix}^{-1} \cap G_1(\mathbb{Q}).$$

Choosing a complete representative set $\{c\} \subset F_{\mathbb{A}}^\times$ for the strict ray class group Cl_F^+ of F, we find by the approximation theorem that

$$G(\mathbb{A}) = \coprod_{c \in Cl_F^+} G(\mathbb{Q}) \left(\begin{smallmatrix} c & 0 \\ 0 & 1 \end{smallmatrix}\right) S_1(\mathfrak{N}) \cdot G(\mathbb{R})^+ = \coprod_{c \in Cl_F^+(\mathfrak{N})} G(\mathbb{Q}) \left(\begin{smallmatrix} c & 0 \\ 0 & 1 \end{smallmatrix}\right) S \cdot G(\mathbb{R})^+,$$

where $G(\mathbb{R})^+$ is the identity-connected component of the Lie group $G(\mathbb{R})$, $Cl_F^+(\mathfrak{N})$ is the strict ray class group modulo $\mathfrak{N}$, and S is either $S_1^1(\mathfrak{N})$ or $S(\mathfrak{N})$. Similarly, by the strong approximation theorem ([LFE] Chapter 9), we have $G_1(\mathbb{A}) = G_1(\mathbb{Q})K \cdot G(\mathbb{R})^+$, where K is given either by

$$G_1(\mathbb{A}^{(\infty)}) \cap \left(\begin{smallmatrix} c & 0 \\ 0 & 1 \end{smallmatrix}\right) S_1^1(\mathfrak{N}) \left(\begin{smallmatrix} c & 0 \\ 0 & 1 \end{smallmatrix}\right)^{-1} \quad \text{or} \quad G_1(\mathbb{A}^{(\infty)}) \cap \left(\begin{smallmatrix} c & 0 \\ 0 & 1 \end{smallmatrix}\right) S(\mathfrak{N}) \left(\begin{smallmatrix} c & 0 \\ 0 & 1 \end{smallmatrix}\right)^{-1}.$$

This shows

$$\begin{aligned}
G(\mathbb{Q})\backslash G(\mathbb{A})/S_1(\mathfrak{N})C_{\mathbf{0}} &\cong \coprod_{\mathfrak{c} \in Cl_F^+} \mathfrak{M}(\mathfrak{c}, \mathfrak{N})(\mathbb{C}), \\
G(\mathbb{Q})\backslash G(\mathbb{A})/S_1(\mathfrak{N})Z(\mathbb{A})C_{\mathbf{0}} &\cong \coprod_{\mathfrak{c} \in Cl_F^+/(Cl_F^+)^2} \mathfrak{M}(\mathfrak{c}, \mathfrak{N})(\mathbb{C}), \\
G(\mathbb{Q})\backslash G(\mathbb{A})/S_1^1(\mathfrak{N})Z(\mathbb{R})C_{\mathbf{0}} &= \coprod_{\mathfrak{c} \in Cl_F^+(\mathfrak{N})} \mathfrak{M}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))(\mathbb{C}), \\
G(\mathbb{Q})\backslash G(\mathbb{A})/S(N)Z(\mathbb{R})C_{\mathbf{0}} &= \coprod_{\mathfrak{c} \in Cl_F^+(\mathfrak{N})} \mathfrak{M}(\mathfrak{c}, \Gamma(N))(\mathbb{C}),
\end{aligned} \tag{4.7}$$

where Z is the center of G and $C_{\mathbf{0}}$ is the stabilizer of $\mathbf{i} = (\sqrt{-1}, \ldots, \sqrt{-1}) \in \mathfrak{Z}$ in $G_1(\mathbb{R})$ (which we regard as the origin $\mathbf{0}$ of $\mathfrak{Z}$).

Recall the identification $X(T) = \mathrm{Hom}_{\text{alg-gp}}(T, \mathbb{G}_m)$ with $\mathbb{Z}[I]$ so that $k(x) = \prod_\sigma \sigma(x)^{k_\sigma}$. Let $\underline{\omega}^k$ over $\mathfrak{M}(\mathfrak{c}, \mathfrak{N})$, $\mathfrak{M}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))$, and $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ be the automorphic vector bundle of weight $k \in X(T) = \mathbb{Z}[I]$ (whose precise definition will be given in Section 4.1.6). Assuming that $F \neq \mathbb{Q}$ (since the case where $F = \mathbb{Q}$ is already treated in Chapter 3), define

$$\begin{aligned}
G_k(\mathfrak{c}, \mathfrak{N}; R) &= H^0(\mathfrak{M}(\mathfrak{c}, \mathfrak{N})_{/R}, \underline{\omega}^k_{/R}), \\
G_k(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}); R) &= H^0(\mathfrak{M}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))_{/R}, \underline{\omega}^k_{/R}), \\
\text{and } G_k(\mathfrak{c}, \Gamma(N); R) &= H^0(\mathfrak{M}(\mathfrak{c}, \Gamma(N))_{/R}, \underline{\omega}^k_{/R})
\end{aligned} \tag{4.8}$$

for $\mathbb{Z}[\frac{1}{d(F)N}]$-algebras R. Regarding $f \in G_k(\mathfrak{c}, \mathfrak{N}; \mathbb{C})$ as a holomorphic function of $z \in \mathfrak{Z}$ by $f(z) = f(\mathcal{L}_z, O_+^\times \lambda_z, i_z)$, it satisfies the following automorphic property by (4.2) (cf. [GME] 2.4.3),

$$f(\gamma(z)) = f(z)(cz + d)^k \quad \text{with } (cz + d)^k = \prod_\sigma (c^\sigma z_\sigma + d^\sigma)^{k_\sigma} \tag{4.9}$$

for all $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_1(\mathfrak{c}, \mathfrak{N})$. We leave it to the reader to formulate the corresponding facts for $f \in G_k(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}); \mathbb{C})$ and $f \in G_k(\mathfrak{c}, \Gamma(N); \mathbb{C})$. The function f has the Fourier expansion $f(z) = \sum_{\xi \in \mathfrak{a}\mathfrak{b}} a(\xi) \mathbf{e}_F(\xi z)$ at the cusp corresponding to $(\mathfrak{a}, \mathfrak{b})$. Here $\mathbf{e}_F(\xi z) = \exp(2\pi\sqrt{-1}\sum_\sigma \xi^\sigma z_\sigma)$. This Fourier expansion gives the q-expansion $f_{\mathfrak{a},\mathfrak{b}}(q)$ substituting q^ξ for $\mathbf{e}_F(\xi z)$ as we show in the following subsection (see also [DAV] Proposition V.1.5).

4.1.4 Toroidal Compactification

We now give an oversimplified description of how to make the toroidal compactification of the Hilbert–Blumenthal moduli space $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ over $\mathbb{Z}[\frac{1}{N}, \mu_N]$, because the main purpose of this chapter is to present a short description of the structure of our theory applied to Hilbert modular varieties and in any case, the process of the toroidal compactification has been well covered in book form by [DAV] and [CSM]. We can formulate the result also for the moduli spaces $\mathfrak{M}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))$ and $\mathfrak{M}(\mathfrak{c}, \mathfrak{N})$ with minor modification which is left to the reader, because in any case these moduli over $\mathcal{W}$ are (geometric) quotients of $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ by a finite constant group action as long as N is prime to p. Main references in addition to the two books are [C], [K3], and [Ra] (and [HT]).

Let $(\mathfrak{a}, \mathfrak{b})$ be a cusp of $\mathfrak{M}(\mathfrak{c}, \Gamma(1))$ for two fractional ideals $\mathfrak{a}$ and $\mathfrak{b}$ (so, $\mathfrak{c} = \mathfrak{a}\mathfrak{b}^{-1}$). Since the cusp is identified with the infinity cusp of $\Gamma(O; \mathfrak{a}, \mathfrak{b})$, it is fixed by $\Gamma_\infty(O; \mathfrak{a}, \mathfrak{b}) = \Gamma(O; \mathfrak{a}, \mathfrak{b}) \cap B(\mathbb{Q})$, where B is the upper triangular Borel subgroup of G. The group $\Gamma_\infty(O; \mathfrak{a}, \mathfrak{b})$ is the stabilizer of the exact sequence $0 \to \mathfrak{a}^* \to (\mathfrak{a}^* \oplus \mathfrak{b}) \to \mathfrak{b} \to 0$ in the sense that $\gamma \in \Gamma(O; \mathfrak{a}, \mathfrak{b})$ acts on the column vector ${}^t(a, b) \in (\mathfrak{a}^* \oplus \mathfrak{b})$ via left multiplication ${}^t(a, b) \mapsto \gamma \cdot {}^t(a, b)$ and $\Gamma_\infty(O; \mathfrak{a}, \mathfrak{b})$ is the stabilizer of $(\mathfrak{a}^* \oplus 0)$. The O-module $H = (\mathfrak{a}^* \oplus \mathfrak{b})$ (column vectors) is identified with $\mathcal{L}_z = \mathfrak{b}z + \mathfrak{a}^* \subset F_{\mathbb{C}}$ by sending ${}^t(a, b)$ to $(a, b)J_1 \cdot {}^t(z, 1) = a - bz$ for $J_1 = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, and the left multiplication by γ^{-1} gives rise to $z \mapsto \gamma(z)$ by $\left((a, b) \cdot {}^t\gamma^{-1}\right) J_1 \cdot {}^t(z, 1) = (a, b)J_1 \cdot {}^t(\gamma(z), 1)(cz + d)$ if $\gamma = \left(\begin{smallmatrix} * & * \\ c & d \end{smallmatrix}\right)$.

Let N be a positive integer prime to p. Since $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ classifies AVRMs with level N-structure $\phi_N : \frac{(N)^*}{O^*} \oplus \frac{O}{N} \cong \frac{(N\mathfrak{a})^*}{\mathfrak{a}^*} \oplus \frac{N^{-1}\mathfrak{b}}{\mathfrak{b}} = \frac{N^{-1}\mathcal{L}_z}{\mathcal{L}_z}$, a cusp of $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ is determined by a triple $(\mathfrak{a}, \mathfrak{b}, \phi_N)$ with $\mathfrak{a}\mathfrak{b}^{-1} = \mathfrak{c}$ modulo the action of $\Gamma_\infty((N); \mathfrak{a}, \mathfrak{b})$. We fix a triple $(\mathfrak{a}, \mathfrak{b}, \phi_N)$ as above and study the toroidal compactification of $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ at the cusp $(\mathfrak{a}, \mathfrak{b}, \phi_N)$, which heavily depends on a chosen cone decomposition of $C = \{\xi \in F_\infty | \xi^\sigma > 0 \text{ for all } \sigma : F \hookrightarrow \mathbb{R}\}$ (the cone of totally positive numbers in $F_\infty = F \otimes_{\mathbb{Q}} \mathbb{R} = \mathbb{R}^I$).

A (simplicial) cone σ in C of dimension $m \le [F : \mathbb{Q}] = g$ is an open span $\sigma = \mathbb{R}_+ v_1 + \mathbb{R}_+ v_2 + \cdots + \mathbb{R}_+ v_m$ inside C for linearly independent $v_1, \ldots, v_m \in C$, and a *cone decomposition* $\mathcal{C}$ of C is a decomposition of C into a disjoint union $\bigsqcup_{\sigma \in \mathcal{C}} \sigma$ of open cones σ of dimension $\le g$. Choose a cone decomposition $\mathcal{C} = \mathcal{C}(\mathfrak{a}, \mathfrak{b}, \phi_N)$ of $C = \bigsqcup_{\sigma \in \mathcal{C}} \sigma$ such that

(PC1) σ is an open simplicial cone;
(PC2) The cones in $\mathcal{C}$ are permuted under multiplication by ϵ^2 for $\epsilon \in T(\mathbb{Z})(N)$, where $T(\mathbb{Z})(N) = \{\epsilon \in O^\times | \epsilon \equiv 1 \mod N\}$. There are finitely many cones modulo $T(\mathbb{Z})(N)$, and $\epsilon(\sigma) \cap \sigma \neq \emptyset$ implies that $\epsilon = 1$;
(PC3) σ is smooth (i.e., generated by a part of a $\mathbb{Z}$-base of $(\mathfrak{a}\mathfrak{b})^*$);
(PC4) $\{\sigma\}$ is sufficiently fine so that the toroidal compactification is projective (see [DAV] IV.2.4 for an exact condition for projectivity).

The existence of such a cone decomposition was remarked by T. Shintani in 1976 (for a proof, see [LAP] IV.5.2 or [LFE] Theorem 2.7.1) and previously also by algebraic geometers studying toroidal compactification ([TEB]). We make a choice *once and for all* of such a cone decomposition $\mathcal{C}_{(\mathfrak{a},\mathfrak{b})}$ for $N = 1$ and use the same decomposition for the cusp $s = (\mathfrak{a}, \mathfrak{b}, \phi_N)$ for every ϕ_N and all $N > 1$. Then we have an action of $\Gamma(O; \mathfrak{a}, \mathfrak{b})$ on $\{\mathcal{C}_s\}_s$ (indexed by equivalence classes s of cusps), and $\Gamma(O; \mathfrak{a}, \mathfrak{b})$ permutes them by $\phi_N \mapsto \gamma \circ \phi_N$.

If $\mathfrak{a}$ and $\mathfrak{b}$ are prime to N, we have a canonical isomorphism

$$\Gamma(O; \mathfrak{a}, \mathfrak{b})/\Gamma(N; \mathfrak{a}, \mathfrak{b}) \cong \Gamma(O; O, \mathfrak{c}^{-1})/\Gamma(N; O, \mathfrak{c}^{-1}),$$

which in turn is isomorphic to $SL_2(O/(N))$ (possibly noncanonically) by conjugation inside $G(\mathbb{A}^{(\infty)})$. Thus we may regard the action on the cusp as an action of $SL_2(O/(N))$.

Each Hilbert modular form f has a q-expansion

$$f(q) = a(0, f) + \sum_{\xi \in N^{-1}\mathfrak{a}\mathfrak{b} \cap C} a(\xi; f) q^{\xi}$$

at the cusp $(\mathfrak{a}, \mathfrak{b}, \phi_N)$. The minimal compactification $M^*(\mathfrak{c}, \Gamma(N))$ of the open modulus space $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ if it exists is characterized by the property that it is covered by any smooth compactification of $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ (having a divisor of normal crossing at the cusps) so that the covering map induces an isomorphism in the interior $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$. Thus we convince ourselves that the formal stalk of the minimal compactification at the cusp $s = (\mathfrak{a}, \mathfrak{b}, \phi_N)$ is given by $H^0(T(\mathbb{Z})(N), \widehat{R}_s(N))$, where

$$\widehat{R}_s(N) = \left\{ a_0 + \sum_{\xi \in N^{-1}\mathfrak{a}\mathfrak{b} \cap C} a_\xi q^\xi \middle| a_\xi \in \mathbb{Z}[\frac{1}{N}, \mu_N] \right\}$$

on which $\epsilon \in T(\mathbb{Z})(N)$ acts by

$$a_0 + \sum_{\xi} a_\xi q^\xi \mapsto a_0 + \sum_{\xi} a_\xi q^{\epsilon^2 \xi}.$$

Here the series $a_0 + \sum_\xi a_\xi q^\xi$ is a formal series whose coefficients a_ξ could be nonzero for all $\xi \in (C \cap N^{-1}\mathfrak{a}\mathfrak{b}) \cup \{0\}$. Thus $\widehat{R}_s(N)$ is the completion of the monoid ring $R_s(N)$ of the semi-group $(N^{-1}\mathfrak{a}\mathfrak{b}) \cap C$ under the adic topology of the augmentation ideal.

Let $\sigma^\vee$ be the dual cone $\sigma^\vee = \{x \in F_\infty | \mathrm{Tr}_{F/\mathbb{Q}}(x\sigma) \geq 0\}$. Then $C = \bigcap_\sigma \sigma^\vee$. To make our conviction feasible, we try to describe the complicated ring $\widehat{R}_s(N)$ using more reasonable rings. We consider the monoid ring $R_\sigma(N)$ of $\sigma^\vee \cap N^{-1}\mathfrak{a}\mathfrak{b}$. Thus

$$R_\sigma(N) = \left\{ a_0 + \sum_{\xi} a_\xi q^\xi \middle| a_\xi \in \mathbb{Z}[\frac{1}{N}, \mu_N],\ \xi \in \sigma^\vee \cap N^{-1}\mathfrak{a}\mathfrak{b} \right\},$$

but the sum $a_0 + \sum_\xi a_\xi q^\xi$ is supposed to be *finite* (so, $a_\xi = 0$ for almost all ξ). For each σ as above, by (PC3), each cone σ is generated by a part of a $\mathbb{Z}$-basis of $N(\mathfrak{ab})^*$, and hence $\sigma^\vee$ is generated by a part of a $\mathbb{Z}$-basis of $N^{-1}\mathfrak{ab}$. Thus we may assume that $(N(\mathfrak{ab})^*) \cap \sigma$ is generated over $\mathbb{Z}$ by $t_1, \ldots, t_r$ $(0 < r \leq [F : \mathbb{Q}] = g)$. We have a base $\xi_1, \ldots, \xi_g$ of $N^{-1}\mathfrak{ab}$ so that $\mathrm{Tr}(t_i \xi_j) = \delta_{ij}$ for $1 \leq i \leq j \leq r$ and $\mathrm{Tr}(t_i \xi_{r+k}) = 0$ for $k > 0$, and each $\xi \in N^{-1}\mathfrak{ab} \cap \sigma^\vee$ can be uniquely written as $\xi = \sum_i m_i \xi_i$ with $m_j \in \mathbb{Z}$ and $m_j \geq 0$ if $j \leq r$. Writing $T_j = q^{\xi_j}$, we find

$$R_\sigma(N) = \mathbb{Z}[\frac{1}{N}, \mu_N][T_1, \ldots, T_g][\frac{1}{T_{r+1}}, \ldots, \frac{1}{T_g}]$$

with $Spec(R_\sigma(N)) = \mathbb{G}_a^r \times \mathbb{G}_m^{g-r}$.

The additive group $\mathbb{G}_a = Spec(\mathbb{Z}[\frac{1}{N}, \mu_N][T])$ is a "local" compactification of $\mathbb{G}_m$ at the origin filling the point "0" onto the hole at the origin of $\mathbb{G}_m$: $\mathbb{G}_a = \mathbb{G}_m \cup \{0\}$, and therefore $Spec(R_\sigma(N))$ is a partial "local" compactification of $Spec(R_\tau(N)) \cong \mathbb{G}_a^{r-1} \times \mathbb{G}_m^{g-r+1}$ for each face τ of σ. Thus we can glue $\{Spec(R_\sigma(N))\}_\sigma$ over the monoid ring $R_s(N)$ to get the toroidal compactification $\mathcal{X}$ of $\{Spec(R_\sigma(N))\}_\sigma$ on which $T(\mathbb{Z})(N)$ acts by translation.

If one adds the origin 0 to $\mathcal{C}$, $0^\vee$ is the total space F_∞, and the corresponding ring is given by the group ring $R_0(N)$ of $N^{-1}\mathfrak{ab}$ isomorphic to

$$\mathbb{Z}\left[\frac{1}{N}, \mu_N\right][T_1, \ldots, T_g][T_1^{-1}, \ldots, T_g^{-1}],$$

and $Spec(R_0(N)) \cong \mathbb{G}_m \otimes (N^{-1}\mathfrak{ab})$ (on which $T(N)(\mathbb{Z})$ acts naturally by $q^\xi \mapsto q^{\epsilon^2 \xi}$). We have the inclusion $Spec(R_0(N)) \subset \mathcal{X}$ equivariant under the action of $T(N)(\mathbb{Z})$. The divisor $\infty = \mathcal{X} - Spec(R_0(N))$ is stable under the action of $T(\mathbb{Z})(N)$. Write $\widehat{\mathcal{X}}$ for the formal completion of $\mathcal{X}$ along ∞, and let $Spf(\widehat{R}_\sigma(N))$ be the formal completion of $Spec(R_\sigma(N)) \subset \mathcal{X}$ along ∞.

We can make a quotient $\mathfrak{X}(N) = \widehat{\mathcal{X}}/T(\mathbb{Z})(N)$. By the work of Mumford [Mu1], $\mathfrak{X}$ carries a semi-abelian scheme $\mathrm{Tate}_s(q)$ having real multiplication (which gives rise to an AVRM over $\mathfrak{X}^\circ(N) := \mathfrak{X}(N) - D_s$ for the image D_s of ∞ in $\mathfrak{X}(N)$) and equipped with a level $\Gamma(N)$-structure and a polarization determined by the cusp s (see Section 6.1.6 for the definition of semi-abelian schemes). Indeed, the tautological homomorphism $q : \mathfrak{ab} \to \mathbb{G}_m(R_0(1))$ sending ξ to q^ξ induces by duality $q : \mathfrak{b} \to (\mathbb{G}_m \otimes \mathfrak{a}^*)(R_0(1))$. We write $q^\mathfrak{b}$ for its image. As was done by Mumford [Mu1] (reproduced in [DAV] in Appendix), we have a semi-abelian scheme $\mathrm{Tate}_s(q)$ over $\mathfrak{X}(N)$ that coincides with an appropriate "quotient": $(\mathbb{G}_m \otimes \mathfrak{a}^*)/q^\mathfrak{b}$ on $\mathfrak{X}(N)$. The quotient is in turn isomorphic to the pullback to $\mathfrak{X}^\circ(N)$ of the universal abelian scheme over $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$. Then by the universality of $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$, D_s gives a toroidal boundary of $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ at s. Performing this process for all cusps s, we obtain the smooth toroidal compactification $M(\mathfrak{c}, \Gamma(N)) = M_\mathcal{C}(\mathfrak{c}, \Gamma(N))$ of $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$.

This Tate AVRM $\mathrm{Tate}_s(q)$ is revisited in the following subsection. Thus by this construction, we have a semi-abelian scheme $\mathcal{G} = \mathcal{G}_{\mathcal{C}}$ over $M_{\mathcal{C}}(\mathfrak{c}, \Gamma(N))$ extending the universal abelian scheme over $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ which induces the Tate AVRM $\mathrm{Tate}_s(q)$ at the cusp $s = (\mathfrak{a}, \mathfrak{b}, \phi_N)$. Here is a reproduction of a statement in [C].

Theorem 4.6 *Let $\mathcal{C} = \{\mathcal{C}_s\}_s$ be the collection of the cone decompositions satisfying* (PC1–4) *for each equivalence class s of cusps of $\mathfrak{M}(\mathfrak{c}, \Gamma(1))$. Then there exists a smooth projective toroidal compactification $M_{\mathcal{C}}(\mathfrak{c}, \Gamma(N))$ dependent on $\mathcal{C}$ such that*

(1) *The semi-abelian scheme $\mathcal{G}_{\mathcal{C}}$ extends the universal abelian scheme over $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ which degenerates into $\mathbb{G}_m \otimes \mathfrak{a}^*$ over the cuspidal divisor D_s for $s = (\mathfrak{a}, \mathfrak{b}, \phi_N)$ and coincides with $\mathrm{Tate}_s(q)$ on the formal completion $\widehat{M}_s(\mathfrak{c}, \Gamma(N))$ of $M_{\mathcal{C}}(\mathfrak{c}, \Gamma(N))$ along D_s.*
(2) *The formal completion $\widehat{M}_s(\mathfrak{c}, \Gamma(N))$ along D_s is isomorphic to $\mathfrak{X}(N)$ over $\mathbb{Z}\left[\frac{1}{N}, \mu_N\right]$.*

We write $\mathfrak{M}$ for one of the moduli spaces $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$, $\mathfrak{M}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))$ and $\mathfrak{M}(\mathfrak{c}, \mathfrak{N})$ defined over $\mathbb{Z}[\frac{1}{N}]$. From the data $\mathcal{C}$, we get a unique toroidal compactification $M_{/\mathbb{Z}[\frac{1}{N}, \mu_N]}$ of $\mathfrak{M}_{/\mathbb{Z}[\frac{1}{N}, \mu_N]}$, which carries a semi-AVRM $\mathcal{G}$ with a level structure (see [C] 3.5.1) canonically determined by the toroidal compactification data.

Once the smooth toroidal compactification M and the universal semi-abelian scheme $\mathcal{G} \xrightarrow{\pi} M$ with respect to $\mathcal{C}$ are constructed, we get a vector bundle $\underline{\omega} = \pi_* \Omega_{\mathcal{G}/M}$ which is locally free of rank $g = [F : \mathbb{Q}]$ over M. The Koecher principle (which can be proven in exactly the same way for $\mathfrak{M} \subset M$ as in [DAV] Proposition V.1.5) tells us

$$H^0(M, \det(\underline{\omega})^{\otimes j}) = H^0(\mathfrak{M}, \det(\underline{\omega})^{\otimes j}) \tag{4.10}$$

for all $j \in \mathbb{Z}$ if $F \neq \mathbb{Q}$ (the case of $F = \mathbb{Q}$ is already treated). Therefore $H^0(M, \det(\underline{\omega})^{\otimes k})$ is independent of the compactification M; so, M covers $M^* = Proj(\mathcal{O}_{\mathcal{M}})$ for $\mathcal{O}_{\mathcal{M}} = \bigoplus_{k \geq 0} H^0(M, \det(\underline{\omega})^{\otimes k})$. Since the theta constants give the modular functions classifying the abelian part $\mathfrak{M}$ (cf. [CSM] and [DAV] Theorem V.2.3), the scheme $M^* = Proj(\mathcal{O}_{\mathcal{M}})$ gives the minimal compactification of $\mathfrak{M}$ adding one point at each cusp, $M^* = \mathfrak{M} \cup \{\text{cusps}\}$. In particular, M^* has the desired formal stalk $\widehat{R}_s(N)$ at each cusp s.

We now study the geometric inertia group of the cuspidal prime divisor. Since characteristic 0 theory is easier, we describe a characteristic p tame inertia group. The prime p is assumed to be unramified in $F/\mathbb{Q}$; so, it is prime to the discriminant $d(F)$ of F. We assume that $\mathcal{W}$ is the strict Henselization of $\mathbb{Z}_{(p)}$ with respect to the fixed embedding $i_p : \overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_p$. Thus $\mathcal{W}$ is the pullback by i_p of the valuation ring of the maximal subfield of $\overline{\mathbb{Q}}_p$ unramified over $\mathbb{Q}_p$. Write $\mathbb{F}$ for the residue field of $\mathcal{W}$, which is an algebraic closure of $\mathbb{F}_p$. Hereafter in this subsection, M_N for N prime to p denotes the special

modulo $\mathfrak{m}$ fiber $M(\mathfrak{c}, \Gamma(N)) \otimes_{\mathbb{Z}[\mu_N]} \mathbb{F}$ for the maximal ideal $\mathfrak{m} \subset \mathcal{W}$. By the projectivity and smoothness of M_N over $\mathbb{Z}\left[\frac{1}{Nd(F)}, \mu_N\right]$, $M_{N/\mathbb{F}}$ is irreducible (Zariski's connectedness theorem [ALG] III.11.3).

We use the same cone decomposition $\mathcal{C}$ for all N. Then we get a finite covering, $M_{N'} := M_{\mathcal{C}}(\mathfrak{c}, \Gamma(N')) \twoheadrightarrow M_{\mathcal{C}}(\mathfrak{c}, \Gamma(N))$ for all multiples N' of N. As long as N' is prime to p, this covering is tamely ramified at the cuspidal divisors $\sum_s D_s$ (over $\mathbb{Z}_{(p)}$), and the closure $\widehat{\Gamma}(O; \mathfrak{a}, \mathfrak{b})^{(p)}$ in $G_1(\mathbb{A}^{(p\infty)})$ naturally acts on $M_{N'}$. Taking the limit $M_{\mathcal{C}}^{(p)} = \varprojlim_{p \nmid N'} M_{N'}$ (N' running over multiples of N prime to p), we have $\mathrm{Gal}(M^{(p)}/M_{N/\mathbb{F}}) = \widehat{\Gamma}(N; \mathfrak{a}, \mathfrak{b})^{(p)}$. However, the full group $G_1(\mathbb{A}^{(p\infty)})$ may not act on $M^{(p)}$.

Let $O^{\times}_{(p)+} = O^{\times}_{(p)} \cap F^{\times}_{+}$, where $O_{(p)} = O \otimes_{\mathbb{Z}} \mathbb{Z}_{(p)}$. Among simplicial cones in $\mathcal{C}/T(\mathbb{Z})(N)$, r-dimensional faces are in bijection with generic points of the $(g-r)$th stratum. This is due to the following reason (see [DAV] IV.2.5). If a cone σ is spanned by n linearly independent elements $w_1, \ldots, w_n$ in $F^{\times}_{+}$, its dual cone $\sigma^{\vee}$ is given by $\mathbb{Q}v_1 + \cdots + \mathbb{Q}v_{g-n} + \mathbb{Q}_+ w_1 + \cdots + \mathbb{Q}_+ w_n$ with $\mathrm{Tr}_{F/\mathbb{Q}}(v_i w_j) = 0$. Note that

$$N(\mathfrak{ab})^* = \left\{ \alpha \in F \middle| \left(\begin{smallmatrix} 1 & \alpha \\ 0 & 1 \end{smallmatrix}\right) \in \widehat{\Gamma}(N; \mathfrak{a}, \mathfrak{b}) \right\}.$$

By (PC3), the choices of $v_1, \ldots, v_{g-n}$ and $w_1, \ldots, w_n$ can be made to be generators over $\mathbb{Z}$ of $N^{-1}(\mathfrak{ab})$. This point that w_j is a part of a $\mathbb{Z}$-base of $N^{-1}(\mathfrak{ab})$ assures us of the smoothness of M_N. By this fact, w_j is prime to p if $p \nmid N$; in other words, $w_j \in O^{\times}_{(p)+}$.

We have $R_\sigma(N) = \mathbb{Z}\left[\frac{1}{Nd(F)}, \mu_N\right][T_i, S_j][T_i^{-1}]$ with $T_i = q^{v_i}$ and $S_j = q^{w_j}$. Thus if we shrink $\widehat{\Gamma}(N; \mathfrak{a}, \mathfrak{b})$ sending $N \to \infty$ (keeping $p \nmid N$), the tower of rings $\{R_\sigma(N)\}_N$ tamely ramifies maximally along the irreducible closed subscheme defined by the set of n-equations $\{S_1 = 0, \ldots, S_n = 0\}$. We find $R_\sigma(N) \otimes_{\mathbb{Z}} \mathcal{W} = \mathcal{W}[T_i, S_j][T_i^{-1}]_{i,j}$; so, this gives rise to an isomorphism

$$\mathbb{G}_m^g \subset Spec(R_\sigma(N)) \cong \mathbb{G}_m^{g-n} \times \mathbb{G}_a^n \subset \mathbb{G}_a^g \quad \text{with } g = [F : \mathbb{Q}].$$

In the process of compactification, étale locally, the hole at 0 in $\mathbb{G}_m^g$ is filled by a closed $(g-n)$-dimensional subscheme $\mathbb{G}_m^{g-n} \times 0$ to get the "local" compactification $Spec(R_\sigma(N))$. By the natural morphism $Spec(R_{\sigma'}(N)) \hookrightarrow Spec(R_\sigma(N))$ for a face σ' of σ, we glue schemes $\{Spec(R_\sigma(N))\}_{\sigma \in \mathcal{C}}$ and then make a formal quotient under the action of $T(\mathbb{Z})(N)$. The formal quotient is then algebraized. The algebraization requires us to impose the conditions (PC1–4) on $\mathcal{C}$ ([DAV] IV.2). Thus g-dimensional cone σ gives a closed point in the highest stratum, and one-dimensional simplicial cones in the closure of σ are in bijection with prime divisors containing the point. By construction, these prime divisors cross each other normally giving the divisor D_s of normal crossing (see [TFG] 1.8 for divisors with normal crossing). Thus the intersection where $n-1$ distinct prime divisors in D_s meet is of dimension $g-n$ which forms the nth stratum of D_s.

When σ is one-dimensional, the corresponding divisor is defined by $S_1 = q^{w_1} = 0$ étale locally; so, its maximal tamely ramified extension is generated by $S_1^{1/n_1} = q^{w_1/n_1}$ (étale locally) for all n_1 prime to p.

Take a generic point η of the cuspidal divisor D_s, which therefore corresponds to a half line $\mathbb{Q}_+^\times \xi_\eta \subset F_+^\times$ with $\xi_\eta \in O_{(p)+}^\times$, where $O_{(p)+}^\times = O_{(p)}^\times \cap F_+^\times$ and $O_{(p)} = O \otimes_{\mathbb{Z}} \mathbb{Z}_{(p)}$. As seen above, $M^{(p)}/M_N$ infinitely ramifies over M_N at a generic point η of an irreducible component $\overline{\{\eta\}}$ of the cuspidal divisor, and by the q-expansion principle, the inertia group $I_{N,\eta}$ of η in $\mathrm{Gal}(M^{(p)}/M_N) = \widehat{\Gamma}(N;\mathfrak{a},\mathfrak{b})^{(p)}$ is given by

$$I_{N,\eta} = \left\{ \sigma = \left(\begin{smallmatrix} 1 & u\xi_\eta \\ 0 & 1 \end{smallmatrix}\right) \in \widehat{\Gamma}(N;\mathfrak{a},\mathfrak{b}) \middle| u \in \mathbb{A}^{(p\infty)} \right\}.$$

More generally, a generic point η_n of the nth stratum is determined by a n-dimensional simplicial cone σ_n. This cone determines a unique submodule $N(\mathfrak{ab})_\eta^*$ of rank n inside $N(\mathfrak{ab})^*$ generated by $\sigma_n \cap N(\mathfrak{ab})^*$ over $\mathbb{Z}$, and the inertia group in $\mathrm{Gal}(M^{(p)}/M_N)$ for η_n is given by

$$I_{N,\eta_n} = \left\{ \sigma = \left(\begin{smallmatrix} 1 & u \\ 0 & 1 \end{smallmatrix}\right) \middle| u \in N(\mathfrak{ab})_\eta^* \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}^{(p)} \right\}.$$

The inertia group is a module under conjugation over $D_{N,\eta_n}/I_{N,\eta_n}$ for the decomposition subgroup $D_{N,\eta_n} \subset \mathrm{Gal}(M^{(p)}/M_N)$ of η_n. If we incorporate the data of this Galois action, we have

Theorem 4.7 *Let the notation and the assumption be as above. In particular, p is a prime unramified in $F/\mathbb{Q}$. Let η_n be a generic point of the nth stratum of the cuspidal divisor D_s of $M_{N/\mathbb{F}}$. Then the p-tame inertia group $I_{N,\eta_n} \subset \mathrm{Gal}(M^{(p)}/M_N)$ of η_n is given by $I_{N,\eta_n} \cong N(\mathfrak{ab})_\eta^* \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}^{(p)}(1)$, as modules over the decomposition group D_{N,η_n}, where $\widehat{\mathbb{Z}}^{(p)}(1) = \varprojlim_{p\nmid N} \mu_N$ as a Galois module.*

The Galois module structure of the inertia group as in the theorem follows from Kummer's theory described in [TFG] Section 2.

4.1.5 Tate Semi-Abelian Schemes with Real Multiplication

Consider the ring $R[[(\mathfrak{ab})_+]]$ $\left((\mathfrak{ab})_+ = \mathfrak{ab} \cap F_+^\times\right)$ made up of all formal series:

$$a_0 + \sum_{\xi \in (\mathfrak{ab})_+} a_\xi q^\xi \quad \text{with } a_\xi \in R$$

for a given base ring R. Take a subset $q^{\mathfrak{ab}_+} = \{q^\xi | \xi \in (\mathfrak{ab})_+\}$ of $R[[(\mathfrak{ab})_+]]$ stable under multiplication, and write $R\{\mathfrak{ab}\}$ for the localization of $R[[(\mathfrak{ab})_+]]$ by the set $q^{\mathfrak{ab}_+}$. Then $\mathbb{Z}[\frac{1}{N}, \mu_N]\{N^{-1}\mathfrak{ab}\}$ is the localization of the formal stalk $\widehat{R}_s(N)$ at the cusp $s = (\mathfrak{a}, \mathfrak{b}, \phi_N)$ of the minimal compactification M_N^*.

The formal completion of the semi-abelian scheme $\mathcal{G}_{\mathcal{C}}$ in Theorem 4.6 along the cuspidal divisors corresponding to the cusp $(\mathfrak{a}, \mathfrak{b})$ is given by the formal quotient $(\widehat{\mathbb{G}}_m \otimes \mathfrak{a}^*)/q^{\mathfrak{b}}$ $(\mathfrak{a}^* = \mathfrak{a}^{-1}\mathfrak{d}^{-1})$, where the inclusion $\xi \mapsto q^{\xi}$ of the additive group $\mathfrak{ab}$ into the multiplicative group $\widehat{\mathbb{G}}_m(\mathbb{Z}[\frac{1}{N}, \mu_N]\{\mathfrak{ab}\})$ induces $q^{\mathfrak{b}} = \{q^{\xi} | \xi \in \mathfrak{b}\} \hookrightarrow (\widehat{\mathbb{G}}_m \otimes \mathfrak{a}^*)(\mathbb{Z}[\frac{1}{N}, \mu_N]\{\mathfrak{ab}\})$ by the isomorphisms:

$$\begin{aligned}
&\mathrm{Hom}(\mathfrak{ab}, \mathbb{G}_m(\mathbb{Z}[\frac{1}{N}, \mu_N]\{\mathfrak{ab}\})) \\
&\qquad \cong \mathrm{Hom}_O(\mathfrak{b}, \mathrm{Hom}(\mathfrak{a}, \mathbb{G}_m(\mathbb{Z}[\frac{1}{N}, \mu_N]\{\mathfrak{ab}\}))) \\
&\qquad\qquad \cong \mathrm{Hom}_O(\mathfrak{b}, (\mathbb{G}_m \otimes \mathfrak{a}^*)(\mathbb{Z}[\frac{1}{N}, \mu_N]\{\mathfrak{ab}\}))
\end{aligned}$$

via the identification $\mathrm{Hom}(\mathfrak{a}, \mathbb{Z}) = \mathfrak{a}^*$ under the trace pairing on F. This formal quotient can be algebraized into a semi-abelian scheme by a work of Mumford [Mu1] (which is included in [DAV] as Appendix). Strictly speaking, by the inclusion $\widehat{R}_s(N) \subset \widehat{R}_\sigma(N)$ for the rings introduced in the previous subsection, over each $Spf(\widehat{R}_\sigma(\mathfrak{ab}))$, we construct the quotient $(\widehat{\mathbb{G}}_m \otimes \mathfrak{a}^*)/q^{\mathfrak{b}}$ and then glue them together to get the formal semi-abelian scheme $\widehat{\mathcal{G}}$ over $\mathfrak{X}(N)$. We denote the algebraization of this formal quotient $(\widehat{\mathbb{G}}_m \otimes \mathfrak{a}^*)/q^{\mathfrak{b}}$ by $\mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q)$ (which is defined over $\mathbb{Z}[\frac{1}{N}, \mu_N]\{\mathfrak{ab}\}$) and call it the *Tate semi-abelian scheme* at the cusp $(\mathfrak{a}, \mathfrak{b})$. The scheme $Spec(\mathbb{Z}[\frac{1}{N}, \mu_N]\{\mathfrak{ab}\})$ is obtained by removing the cusp s from its formal neighborhood $Spf(\widehat{R}_s(N))$ of $s \in M_N^*$. The semi-AVRM coincides with $\mathcal{A} \times_{\mathfrak{M}} Spec(\mathbb{Z}[\frac{1}{N}, \mu_N]\{\mathfrak{ab}\})$ which is the universal abelian scheme $\mathcal{A}$ over $\mathfrak{M}$ restricted to $Spec(\mathbb{Z}[\frac{1}{N}, \mu_N]\{\mathfrak{ab}\})$, and over the cuspidal divisor D_s, the connected component of $\mathcal{G}$ is isomorphic to $\mathbb{G}_m \otimes \mathfrak{a}^*$.

If $\varpi \in O$ is prime to $\mathfrak{a}$, we have a canonical exact sequence $0 \to \mathfrak{a}^{-1} \xrightarrow{\varpi} \mathfrak{a}^{-1} \to O/(\varpi) \to 0$. Tensoring this sequence with $\mathbb{G}_m \otimes \mathfrak{d}^{-1}$ over O, we get another exact sequence:

$$0 \to \mathrm{Tor}_1^O(\mathbb{G}_m \otimes \mathfrak{d}^{-1}, O/(\varpi)) \to \mathbb{G}_m \otimes \mathfrak{a}^* \xrightarrow{\varpi} \mathbb{G}_m \otimes \mathfrak{a}^* \to 0,$$

since $\mathbb{G}_m \otimes \mathfrak{d}^{-1}$ is divisible and $\mathbb{G}_m \otimes \mathfrak{d}^{-1} \otimes_O \mathfrak{a}^{-1} = \mathbb{G}_m \otimes \mathfrak{a}^*$. Since O is a Dedekind domain, we have $\mathrm{Tor}_1^O(\mathbb{G}_m \otimes \mathfrak{d}^{-1}, O/(\varpi)) \cong O/\varpi$, and the above exact sequence identifies $\mathbb{G}_m \otimes \mathfrak{a}^*[\varpi]$ with $\mu_{(\varpi)}$. If $\mathfrak{N}$ is prime to $\mathfrak{a}$, taking $\varpi \in \mathfrak{N}$ prime to $\mathfrak{a}$, we find $\mu_{\mathfrak{N}} \subset \mu_{(\varpi)} \subset \mathbb{G}_m \otimes \mathfrak{a}^*$. This inclusion combined with the projection $\widehat{\mathbb{G}}_m \otimes \mathfrak{a}^* \twoheadrightarrow \mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q)$ gives us a canonical level $\Gamma_1(\mathfrak{N})$-structure $i_{can,\mathfrak{N}} : \mu_{\mathfrak{N}} \hookrightarrow \mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q)$ as long as $\mathfrak{a}$ is prime to $\mathfrak{N}$. If further, $\mathfrak{b}$ is prime to $\mathfrak{N} = (N)$, $N^{-1}\mathfrak{b}/\mathfrak{b}$ is canonically isomorphic to N^{-1}/O; so, we get a level $\Gamma(N)$-structure sending $(\zeta, \xi) \in (\mu_N \otimes O^*) \times (N)^{-1}/O$ to $i_{can,(N)}(\zeta)q^{\xi} \in \mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q)$ which we write as

$$\phi_{can,N} : (\mu_N \otimes O^*) \times (N)^{-1}/O \cong \mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q)[N]$$

as long as N is prime to $\mathfrak{a}$ and $\mathfrak{b}$. This level structure is defined over $\widehat{R}_s(N)$, although $\mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q)$ itself is defined over the smaller ring $\widehat{R}_s(1)$.

For a positive integer n, we have a perfect pairing of group schemes

$$\mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q)[n] \times \mathrm{Tate}_{\mathfrak{b},\mathfrak{a}}(q)[n] \to \mu_n$$

given as follows: Taking $x \in \mathbb{G}_m \otimes \mathfrak{a}^*$ and $y \in \mathbb{G}_m \otimes \mathfrak{b}^*$ such that $x^n = q^\xi$ for $\xi \in \mathfrak{b}$ and $y^n = q^\eta$ for $\eta \in \mathfrak{a}$ and writing $[x] \in \mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q)[n]$ and $[y] \in \mathrm{Tate}_{\mathfrak{b},\mathfrak{a}}(q)[n]$ for the points these x and y represent, we define

$$\langle [x], [y] \rangle_n = x^\eta / y^\xi \in \mu_n,$$

because $(x^\eta/y^\xi)^n = q^{\eta\xi}/q^{\xi\eta} = 1$. Since we may identify the complex points of the Tate AVRM $\mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q)(\mathbb{C})|_{q=\exp(\mathrm{Tr}(z))}$ with $(\mathbb{C}^\times)^I/e(\mathcal{L}_z)$ ($\mathcal{L}_z = 2\pi\sqrt{-1}(\mathfrak{b}z + \mathfrak{a}^*)$) for the variable $z \in \mathfrak{Z}$, where $e : \mathbb{C}^I \to (\mathbb{C}^\times)^I$ given by $(u_\sigma)_\sigma \mapsto (\exp(u_\sigma))$, the polarization λ_z induces the above pairing for any z. Thus the polarization of the universal AVRM over $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ induces an isomorphism $\varphi : {}^t\mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q) \cong \mathrm{Tate}_{\mathfrak{b},\mathfrak{a}}(q)$ so that the Cartier duality pairing ${}^t\mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q)[M] \times \mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q)[M] \to \mu_M$ (in (P1–3) in the proof of Theorem 8.9) composed with φ coincides with the above pairing $\langle \cdot, \cdot \rangle_M$ for all M. Then writing $\mathfrak{c} = \mathfrak{a}\mathfrak{b}^{-1}$, the natural isomorphism

$$\mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q) \otimes \mathfrak{c} = (\mathbb{G}_m \otimes \mathfrak{a}^*/q^{\mathfrak{b}}) \otimes \mathfrak{a}\mathfrak{b}^{-1} = (\mathbb{G}_m \otimes \mathfrak{b}^*)/q^{\mathfrak{a}} = \mathrm{Tate}_{\mathfrak{b},\mathfrak{a}}(q)$$

composed with φ^{-1} gives rise to a canonical $\mathfrak{c}$-polarization λ_{can} on $\mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q)$. Thus by definition, the level structure $\phi_{can,N}$ is compatible with the Weil pairing on $\mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q)[N]$ as required for level $\Gamma(N)$-structures. Similarly, the polarization λ_{can} combined with $i_{can,\mathfrak{N}}$ also gives a $\Gamma_1^1(\mathfrak{N})$-structure (still denoted by $i_{can,\mathfrak{N}}$) of $\mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q)$.

4.1.6 Hasse Invariant and Sheaves of Cusp Forms

Hereafter we often make the base change of our moduli spaces $\mathfrak{M}$, M and M^* of type $\Gamma = \Gamma(N), \Gamma_1^1(\mathfrak{N})$, and $\Gamma_1(\mathfrak{N})$ to the valuation ring $\mathcal{W}$ and regard these schemes as defined over $\mathcal{W}$. Assume that $\mathcal{W} \otimes O = \mathcal{W}^I$ by $w \otimes a \mapsto w\sigma(a)$; so, $\mathcal{W}$ contains all conjugates of O over $\mathbb{Z}$, and p is unramified in $O/\mathbb{Z}$. Further assume that the polarization ideal $\mathfrak{c}$ is prime to the residual characteristic p of $\mathcal{W}$. We fix a cone decomposition $\mathcal{C}$ as in (PC1–4) once and for all and write $\mathcal{G}_{/M}$ for the universal semi-AVRM with respect to $\mathcal{C}$.

A little more generally, we start with a $\mathcal{W}$-scheme S carrying a semi-abelian scheme $\mathcal{G}$ with real multiplication by O. Thus over an open dense subscheme $S^\circ \subset S$, $\mathcal{G} \times_S S^\circ$ is an object of $\mathcal{A}_F$, and over the closed subscheme $Z = S - S^\circ$, $\mathcal{G}$ is a smooth multiplicative group scheme of finite type whose connected component is given by $\mathbb{G}_m \otimes O$ (which is canonically isomorphic to a more standard $\mathbb{G}_m \otimes \mathfrak{d}^{-1}$ because of unramifiedness of p in $F/\mathbb{Q}$). Writing the structure homomorphism of $\mathcal{G}$ as $\pi : \mathcal{G} \to S$, we have a vector bundle

$\underline{\omega} = \pi_* \Omega_{\mathcal{G}/S}$ over S. The pullback action of the endomorphism O makes $\underline{\omega}$ into a schematic $T_{/S}$-module for $T = \mathrm{Res}_{O/\mathbb{Z}} \mathbb{G}_m$. Since p is unramified in $F/\mathbb{Q}$, identifying $X(T)$ with $\mathbb{Z}[I]$, we find $\underline{\omega} = \bigoplus_{\sigma \in I} \underline{\omega}^\sigma$, where $\underline{\omega}^\sigma$ is the σ-eigenspace of T. Then we define $\underline{\omega}^k = \otimes (\underline{\omega}^\sigma)^{\otimes k_\sigma}$ for $k \in \mathbb{Z}[I]$, which is an invertible sheaf over S. The sheaf $\bigoplus_{k \in \mathbb{Z}[I]} \underline{\omega}^k$ of graded $\mathcal{O}_S$-algebras is the affine ring of the scheme $\mathcal{M}_{\mathcal{G}/S} = Spec_S(\mathcal{S})$ (see [GME] 1.5.4) representing the following functor $\mathcal{P}_{\mathcal{G}/S} : S\text{–}SCH \to SETS$ given by

$$\mathcal{P}_{\mathcal{G}/S}(S') = [(A, \omega)_{/S'} \big| A \cong \mathcal{G} \times_S S',\ H^0(A, \Omega_{A/S'}) = (O \otimes_{\mathbb{Z}} \mathcal{O}_{S'})\omega],$$

because having S-morphism $S' \xrightarrow{f} \mathcal{M}$ is equivalent to having an isomorphism $\omega^\sigma \cong \mathcal{O}_{S'}$ for each $\sigma \in I$, which is equivalent to having a generator $\omega \in \underline{\omega}$ over $O \otimes_{\mathbb{Z}} \mathcal{O}_{S'}$. We note that $\det(\underline{\omega})^{\otimes j} = \bigwedge_\sigma (\underline{\omega}^\sigma)^{\otimes j} \cong \underline{\omega}^{jI}$.

We now construct the *Hasse invariant* relative to $\mathcal{G}_{/S}$. Let R be a $\mathcal{W}$-algebra of characteristic p and $(A, \omega) \in \mathcal{P}_{\mathcal{G}/S}(S')$ be a pair of a semi-abelian scheme A over $S' = Spec(R) \to S$ of relative dimension g and a base $\omega = \{\omega_1, \ldots, \omega_g\}$ of $H^0(A, \Omega_{A/S'})$ over R. We have the absolute Frobenius endomorphism $F_{abs} : A_{/R} \to A_{/R}$. Let $\mathcal{T}_{A/S'}$ be the relative tangent bundle; so, $H^0(A, \mathcal{T}_{A/S'})$ is spanned by the dual base $\eta = \eta(\omega)$. For each derivation D of $\mathcal{O}_{A,\mathbf{0}}$, by the Leibniz formula, we have

$$D^p(xy) = \sum_{j=0}^{p} \binom{p}{j} D^{p-j} x D^j y = x D^p y + y D^p x.$$

Thus D^p is again a derivation. The association $D \mapsto D^p$ induces an F_{abs}-linear endomorphism F^* of $\mathcal{T}_{A/S'}$. Then we define $H(A, \omega) \in R$ by $F^* \wedge^g \eta = H(A, \omega) \wedge^g \eta$. Since $\eta(\lambda\omega) = \lambda^{-1}\eta(\omega)$ for $\lambda \in GL_g(R)$, we see

$$\begin{aligned} H(A, \lambda\omega) \wedge^g \eta(\lambda\omega) &= F^*(\det(\lambda)^{-1} \wedge^g \eta(\omega)) = \det(\lambda)^{-p} F^* \wedge^g \eta(\omega) \\ &= \det(\lambda)^{-p} H(A, \omega) \wedge^g \eta(\omega) = \det(\lambda)^{1-p} H(A, \omega) \wedge^g \eta(\lambda\omega). \end{aligned}$$

Thus we get $H(A, \lambda\omega) = \det(\lambda)^{1-p} H(A, \omega)$. We call a semi-abelian scheme A *ordinary* if μ_p^g can be embedded into $A[p]$ étale locally (i.e., after a faithfully flat étale base change, we achieve the closed immersion of group schemes, $\mu_p^g \hookrightarrow A[p]$). In the same manner as in the elliptic curve case (see Section 3.2.6 and [GME] 2.9.1), we know

$$H(A, \omega) = 0 \iff A \text{ is not ordinary.}$$

Now we return to $\mathcal{G}_{/S}$ given by $\mathcal{G}_{/M}$ for the modulus $\mathfrak{M}$ and its toroidal compactification M over $\mathbb{Z}[\frac{1}{Nd(F)}, \mu_N]$. By a result of Moret–Bailly ([DAV] Proposition V.2.1), the graded algebra $\mathcal{O}_{\mathcal{M}} = \bigoplus_{j \geq 0} H^0(M, \underline{\omega}^{jI})$ is of finite type over $\mathbb{Z}[\frac{1}{Nd(F)}, \mu_N]$. Recall the Koecher principle (4.10): $H^0(M, \underline{\omega}^{jI}) = H^0(\mathfrak{M}, \underline{\omega}^{jI})$ for all $j \in \mathbb{Z}$ if $F \neq \mathbb{Q}$ (the case of $F = \mathbb{Q}$ is already treated). Thus

$\mathcal{O}_{\mathcal{M}} = \bigoplus_{\jmath \geq 0} H^0(\mathfrak{M}, \underline{\omega}^{\jmath I})$, and $\underline{\omega}^I$ is ample on $M^* = Proj(\mathcal{O}_{\mathcal{M}})$ (however, $\underline{\omega}^I$ is not ample on M). Then for a sufficiently large integer $a > 0$, we have $H^0(M^*, \underline{\omega}_{/W}^{a(p-1)I}) \otimes_W \mathbb{F} = H^0(M^*_{/\mathbb{F}}, \underline{\omega}_{/\mathbb{F}}^{a(p-1)I})$ for the residue field $\mathbb{F}$ of W (by a theorem of Serre; see [ALG] III Theorem 5.2), and we have a lifting

$$E \in H^0(\mathfrak{M}_{/W}, \underline{\omega}^{a(p-1)I}) = H^0(M^*_{/W}, \underline{\omega}^{a(p-1)I}) \overset{(4.10)}{=} H^0(M_{/W}, \underline{\omega}^{a(p-1)I}) \tag{E}$$

of H^a; that is, $H^a = (E \mod p)$. Since the Hasse invariant is a nontrivial section of $\underline{\omega}^{(p-1)I} = \det(\underline{\omega})^{p-1}$ for $I = \sum_\sigma \sigma$, the scheme $S^* = M^*[\frac{1}{E}]$ is affine and irreducible (because E is a section of an ample line bundle). Then $S^* \subset M^*$ is defined by $S^* = Spec(\mathcal{O}_{\mathcal{M}}/(E-1))$ for $\mathcal{O}_{\mathcal{M}} = \bigoplus_{\jmath \geq 0} H^0(M, \underline{\omega}^{\jmath I})$. Define $S^*_m = S^* \otimes_W W_m$, which is affine, and $S^*_\infty = \varprojlim_m S^*_m$ is an affine formal scheme. Similarly, we put $S = M[\frac{1}{E}]$, and define $S_m = S \otimes_W W_m$. We have a formal scheme $S_\infty = \varprojlim_m S_m$. These (formal) schemes S_m, S_∞, and S are not necessarily affine if $F \neq \mathbb{Q}$.

Let M_∞ be the formal completion of M along the modulo p special fiber $M_1 = M \otimes_W W_1$ for $W_1 = W/pW$. Then $S_\infty \subset M_\infty$ is the ordinary locus of M_∞; that is, S_∞ is the maximal formal subscheme of M_∞ on which the connected component $\mathcal{G}[p]^\circ$ of $\mathcal{G}[p]$ is isomorphic to μ_p^g étale locally. Let

$$\begin{aligned} T_{m,n/W_m} = \mathrm{Isom}_O(O \otimes \mu_{p^n}, \mathcal{G}[p^n]^\circ) &\cong \mathrm{Isom}_O(\widehat{\mathcal{G}[p^n]^\circ}, O/p^nO) \\ &\cong \widehat{\mathcal{G}[p^n]^\circ} - \bigcup_{\mathfrak{n} \supsetneq p^n} \widehat{\mathcal{G}[p^n]^\circ}[\mathfrak{n}]. \end{aligned} \tag{4.11}$$

Then $T_{m,n}/S_m$ is an étale covering with Galois group $T(\mathbb{Z}/p^n\mathbb{Z}) = (O/p^n)^\times$ for $T = \mathrm{Res}_{O/\mathbb{Z}}\mathbb{G}_m$ (called the Hilbert modular Igusa tower over S_m). By a result of K. A. Ribet [Ri], $T_{m,n}$ is irreducible (see Theorem 4.21 for another proof of this fact). By the irreducibility, we get the following q-expansion principle for any p-adically complete W-algebra $R = \varprojlim_m R/p^mR$ (see Corollary 4.23).

(q-exp) *$f(q) = 0$ in $R[[q^\xi]]_\xi \iff f = 0$ over $T_{m,n/R}$ for a section f of the line bundle over $T_{m,n/R}$ defined at ∞.*

The sheaf $\underline{\omega}_{\infty/S_\infty} = \mathcal{O}_{S_\infty} \otimes_{\mathbb{Z}_p} \varprojlim_n \widehat{\mathcal{G}[p^n]^\circ}$ is isomorphic to the dual $\underline{\omega}_{/M}$ of $\pi_* Lie(\mathcal{G}_{/M})$ for $\pi : \mathcal{G} \to M$, because $Lie(\mathcal{G}[p^n]^\circ_{S_n}) = Lie(\mathcal{G}_{/S_m})$ (see Corollary 4.2). In other words, $\underline{\omega}_{/M} = Hom(\pi_* Lie(\mathcal{G}_{/M}), \mathcal{O}_M) = \pi_* \Omega_{\mathcal{G}/M}$ is the algebraization of the formal sheaf $\underline{\omega}_\infty$ on S_∞ (which is uniquely determined by S_1 independently of the choice of E). We define $\underline{\omega}_k \subset \underline{\omega}^k$ by the invertible subsheaf of $\underline{\omega}^k$ made up of sections of $\underline{\omega}^k$ vanishing over the cuspidal divisor $\mathcal{D}_{cusp} = \pi^{-1}(M - \mathfrak{M})$ on M. We call $\underline{\omega}_k$ the *sheaf of cusp forms* of weight k.

4.1.7 p-Adic Hilbert Modular Forms of Level $\Gamma(N)$

Fix a sufficiently fine simplicial cone decomposition $\mathcal{C}$ of C for cusps of $\mathfrak{M}(\mathfrak{c}, \Gamma(1))$ so that we have a smooth projective toroidal compactification

$M = M_{\mathcal{C}}(\mathfrak{c}, \Gamma(N))$ for all $N > 0$ (using this $\mathcal{C}$) and the minimal compactification $M^* = M^*(\mathfrak{c}, \Gamma(N))$ of $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$.

Fix a positive integer N prime to p. We put $S = M[\frac{1}{E}]$ and $S^* = M^*[\frac{1}{E}]$ for a lift E of the Hasse invariant as in (E). Recall $M_m = M \otimes_W W_m$, $S_m = S \otimes_W W_m$, and $S_m^* = S^* \otimes_W W_m$ for $W_m = W/p^m W = \mathcal{W}/p^m\mathcal{W}$. The proof of the vertical control theorem we have given for elliptic modular forms extends to our Hilbert modular case almost intact except for the following hypothesis (see the second identity in (3.2)),

$$\text{(Hp1)} \qquad H^0(S, \underline{\omega}_k) \otimes_W W_m = H^0(S_m, \underline{\omega}_k \otimes_W W_m) \text{ for } S_m = S \otimes_W W_m$$

for the sheaf $\underline{\omega}_k \subset \underline{\omega}^k$ of cusp forms of weight k. In the elliptic modular case, we have proven this resorting to the affineness of the open curve S_m there. In the present case of $F \neq \mathbb{Q}$, the scheme S_m is not affine; so, we need to verify this. This is a subtle point, and *this fact is only valid for the cuspidal sheaf* $\underline{\omega}_k$ *not for* $\underline{\omega}^k$, as we show.

Write $\pi : S \to S^*$. Since by definition $H^0(S, \mathcal{L}) = H^0(S^*, \pi_*(\mathcal{L}))$ for any sheaf $\mathcal{L}_{/S}$ and S^* is affine, we need to verify

$$\pi_*(\underline{\omega}_{k/S}) \otimes_W W_m = \pi_*(\underline{\omega}_{k/S} \otimes_W W_m). \qquad (*)$$

If this holds, we have

$$H^0(S, \underline{\omega}_k) \otimes_W W_m = H^0(S^*, \pi_*(\underline{\omega}_{k/S})) \otimes W_m \overset{(1)}{=} H^0(S^*, \pi_*(\underline{\omega}_{k/S}) \otimes W_m)$$
$$\overset{(*)}{=} H^0(S^*, \pi_*(\underline{\omega}_{k/S} \otimes_W W_m)) = H^0(S_m, \underline{\omega}_k \otimes_W W_m)$$

as desired, where the identity (1) holds by the affineness of S^* (i.e., the ampleness of $\underline{\omega}^I$ on M^*).

We prove $(*)$ stalk by stalk. Outside the cusps, the two sheaves are the same; so, nothing to prove. Recall that for an element $\xi \in F$, we write $\xi \gg 0$ if $\sigma(\xi) > 0$ for all $\sigma \in I$. We have for each cusp x associated with the ideals $\mathfrak{a}\mathfrak{b}^{-1} = \mathfrak{c}$:

$$\widehat{\pi_*(\underline{\omega}^k_{/R})}_x \cong H^0(T(\mathbb{Z})(N), R[[\frac{1}{N}(\mathfrak{a}\mathfrak{b})_{\geq 0}]])$$
$$= \left\{ \sum_{\xi \in \frac{1}{N}(\mathfrak{a}\mathfrak{b})_{\geq 0}} a(\xi) q^\xi \in R[[\frac{1}{N}(\mathfrak{a}\mathfrak{b})_{\geq 0}]] \middle| a(\epsilon^2 \xi) = \epsilon^k a(\xi) \ \forall \epsilon \in T(\mathbb{Z})(N) \right\},$$

where $(\mathfrak{a}\mathfrak{b})_{\geq 0} = \{\xi \in \mathfrak{a}\mathfrak{b} | \xi \gg 0\} \cup \{0\} = C \cap \mathfrak{a}\mathfrak{b}$, and $\epsilon \in T(\mathbb{Z})(N)$ in (PC2) acts on $R[[\frac{1}{N}(\mathfrak{a}\mathfrak{b})_{\geq 0}]]$ by $\epsilon \cdot \left(\sum_{\xi \in (\mathfrak{a}\mathfrak{b})_{\geq 0}} a(\xi) q^\xi\right) = \sum_{\xi \in (\mathfrak{a}\mathfrak{b})_{\geq 0}} \epsilon^{-k} a(\epsilon^2 \xi) q^\xi$. We define $\underline{\omega}_k$ by requiring that its stalk at every cusp be given by those q-expansions vanishing at the cusp. The group cohomology $H^0(U, ?)$ for $U = T(\mathbb{Z})(N)^2$ commutes with taking the tensor product $\otimes_W W_m$ if $H^1(U, ?) = 0$. The vanishing holds if ? is isomorphic to a product of copies of $\mathrm{Hom}(\mathbb{Z}[U], R) = R^U$

(e.g., [CGP] III.6.6). This is the case for any $R[U]$-submodule of $R[[\frac{1}{N}(\mathfrak{ab})_{\geq 0}]]$ made up of q-expansions without the constant term. Then we get

$$\pi_*\widehat{(\underline{\omega}_{k/W})}_x \otimes_W W_m \cong \pi_*\widehat{(\underline{\omega}_{k/W_m})}_x. \tag{4.12}$$

Remark 4.8 If k is a multiple weight congruent modulo $|(O/pO)^\times|$ to a parallel weight $k_0 I$, the character k and $k_0 I$ is identical on $T(\mathbb{F}_p)$, and hence we often have a nontrivial section $E \in H^0(M_1, \underline{\omega}^k)$ congruent modulo p to an Eisenstein series $E_0 \in H^0(M_{/W}, \underline{\omega}^{k_0 I})$ with nontrivial constant term modulo p. Thus $H^0(M_1, \underline{\omega}^k)$ is bigger than $H^0(M_1, \underline{\omega}_k)$. However, by the Koecher principle, $H^0(M_{/W}, \underline{\omega}^k) \otimes_{\mathbb{Z}} \mathbb{Q} = H^0(M_{/W}, \underline{\omega}_k) \otimes_{\mathbb{Z}} \mathbb{Q}$. Thus we conclude that $H^0(M_{/W}, \underline{\omega}^k) \otimes_W W_1 \neq H^0(M_1, \underline{\omega}^k)$ because of W-flatness of $\underline{\omega}^k_{/M}$.

We put

$$\begin{aligned}
V_{m,n} &= H^0(T_{m,n}, \mathcal{O}_{T_{m,n}}), \quad \mathcal{V} = \varinjlim_m V_{m,\infty}, \\
V &= \varprojlim_m V_{m,\infty}, \quad R_\ell = \bigoplus_{k>\ell} H^0(M, \underline{\omega}^k), \\
V^{cusp}_{m,n} &= H^0(T_{m,n}, \mathcal{O}_{T_{m,n}}(-\mathcal{D}_{cusp})), \quad \mathcal{V}_{cusp} = \varinjlim_m V^{cusp}_{m,\infty}, \\
V_{cusp} &= \varprojlim_m V^{cusp}_{m,\infty}, \quad R^{cusp}_\ell = \bigoplus_{k>\ell} H^0(M, \underline{\omega}_k), \\
D_\ell &= \beta(R_\ell)[\frac{1}{p}] \bigcap V, \quad D^{cusp}_\ell = \beta(R^{cusp}_\ell)[\frac{1}{p}] \bigcap V_{cusp},
\end{aligned} \tag{4.13}$$

where $\mathcal{D}_{cusp} = (M - \mathfrak{M}) \times_M T_{m,n}$ is the cuspidal divisor and $\beta(\sum_k f_k) = \sum_k \beta_k(f_k)$ for $f_k \in H^0(M, \underline{\omega}^k)$ with the morphism β_k sending classical modular forms of weight k to p-adic modular forms:

$$\beta_k(f_k)(A, \lambda, \phi_p) = f_k(A, \lambda, \phi_{p,*}\omega_{can})$$

for the canonical differential ω_{can} induced by the $\frac{dt}{t} \otimes 1$ on $\mathbb{G}_m \otimes \mathfrak{d}^{-1}$. Here $k > \ell$ means that $k_\sigma > \ell_\sigma$ for all $\sigma \in I$. As we show in Section 8.1.4 (Corollary 8.4) through a formal argument, the subspece $D^{cusp}_\ell \subset V^{cusp}$ is dense under the p-adic topology of V^{cusp}.

In this $GL(2)$-case, we define in Section 4.1.10 two Hecke operators $U(p)$ (acting on cusp forms of level divisible by p) and $T(p) = T(1,p)$ (acting on cusp forms of a level prime to p) preserving integrality of $\underline{\omega}_k$. The operator $U(p)$ has its effect on q-expansion $a(\xi, f|U(p)) = a(\xi p, f)$ and decreases the level to p^{n-1} from p^n for $n > 1$, and if $k > 2I$, then $T(p) \equiv U(p) \mod p$. Let $e = \lim_{n\to\infty} U(p)^{n!}$ (resp. $e^\circ = \lim_{n\to\infty} T(p)^{n!}$) be the idempotent attached to $U(p)$ (resp. $T(p)$). We attach a subscript or superscript *ord* to the object after applying the idempotent e or e° (depending on the setting). From this, we conclude

Theorem 4.9 *Let the notation be as above; in particular, $M = M(\mathfrak{c}, \Gamma(N))$ and $S = M[\frac{1}{E}]$ for a lift E of the Hasse invariant as in* (E)*. Suppose that $N \geq 3$ and that p is prime to $Nd(F)\mathfrak{c}$. Then we have the following facts.*

(1) *The submodule* D_ℓ^{cusp} *is dense in* V_{cusp};
(2) *The Pontryagin dual* $\mathcal{V}_{cusp}^{ord,*}(=\mathrm{Hom}_W(V_{cusp}^{ord}, W))$ *of* $\mathcal{V}_{cusp}^{ord}$ *is a projective* $W[[T(\mathbb{Z}_p)]]$*-module of finite type;*
(3) $\mathcal{V}_{cusp}^{ord,*} \otimes_{W[[T(\mathbb{Z}_p)]],k} W \cong \mathrm{Hom}_W(H^0_{ord}(S, \underline{\omega}_k), W)$ *if* $k \geq 3I$ $(I = \sum_\sigma \sigma)$, *where the first tensor product is with respect to the algebra homomorphism* $W[[T(\mathbb{Z}_p)]] \to W$ *induced from the character* k *sending* x *to* $\prod_\sigma \sigma(x)^{k_\sigma}$;
(4) *If* $k \geq 3I$, e *induces an isomorphism*

$$H^0_{ord}(S, \underline{\omega}_k) \cong H^0_{ord}(M, \underline{\omega}_k),$$

where $H^0_{ord}(S, ?) = eH^0(S, ?)$ *and* $H^0_{ord}(M, ?) = e^\circ H^0(M, ?)$.

Assertion (3) in the theorem is (the Pontryagin dual version of) the "precise control" described in Section 1.2.

Proof. As we remarked, assertion (1) follows from a more general result which is given in Chapter 8 (see Corollary 8.4) through the theory of false modular forms. We do not prove the assertion (1) here. It is known that

$$\dim H^0_{ord}(M, \underline{\omega}_k) \text{ depends only on } k|_{T(\mathbb{F}_p)} \text{ if } k \geq 3I. \tag{4.14}$$

From this, the assertions (2), (3) and (4) follow by the same argument as in the elliptic modular case (see Section 3.3). The argument giving this exact bound (4.14) is rather involved; so, we allocate the exposition to its own section, Section 4.3 (see Theorem 4.37). Here we just prove the boundedness of $\dim H^0_{ord}(M, \underline{\omega}_k)$ (independent of k) assuming the existence of a Hecke equivariant inclusion map $G_k(\mathfrak{c}, \Gamma; \mathbb{C}) \hookrightarrow H^g(\Gamma, L(k^*; \mathbb{C}))$ (the generalized Eichler–Shimura map for $k^* = k - 2I$; see Theorems 4.36 and 5.13) for the arithmetic subgroup $\Gamma = \Gamma(\mathfrak{c}, N)$ of $SL_2(F) = G_1(\mathbb{Q})$ giving $\mathfrak{M}(\mathfrak{c}, \Gamma(N))(\mathbb{C}) = \Gamma\backslash\mathfrak{Z}$. By the definition (4.3), $\Gamma \subset SL_2(O_{(p)})$ for $O_{(p)} = O_p \cap F$. Here $g = [F : \mathbb{Q}]$, but actually, the following argument is valid for any integer $g \geq 0$. Note that $G_k(\mathfrak{c}, \Gamma; \mathcal{W}) \otimes_{\mathcal{W}} \mathbb{C} = G_k(\mathfrak{c}, \Gamma; \mathbb{C})$ and $G_k(\mathfrak{c}, \Gamma; W) = G_k(\mathfrak{c}, \Gamma; \mathcal{W}) \otimes_{\mathcal{W}} W$ (flat base change; see [GME] Lemma 1.10.2), because these are spaces of global sections of the invertible sheaf $\underline{\omega}^k$. By the long exact sequence attached to the short one: $L(k^*; \mathcal{W}) \xrightarrow{p} L(k^*; \mathcal{W}) \twoheadrightarrow L(k^*; \mathbb{F})$ for $\mathbb{F} = \mathcal{W}/p\mathcal{W}$ ([MFG] Corollary 4.28), we find that $H^g(\Gamma, L(k^*; \mathcal{W})) \otimes_{\mathcal{W}} \mathbb{F}$ injects into $H^g(\Gamma, L(k^*; \mathbb{F}))$; so, we need to bound $\dim_{\mathbb{F}} H^g_{ord}(\Gamma, L(k^*; \mathbb{F}))$. We may identify $L(k^*; R)$ with the space of R-integral polynomials in $(X_\sigma, Y_\sigma)_{\sigma \in I}$ homogeneous of degree $k^*_\sigma = k_\sigma - 2$ for each pair (X_σ, Y_σ). The action of $\gamma \in M_2(O_p)$ on $L(k^*; W)$ is given by $\Phi((X_\sigma, Y_\sigma)^t \sigma(\gamma)^\iota)$, where $xx^\iota = \det(x)$. Let

$$\Gamma' = \{\alpha \in \Gamma | \alpha \bmod pM_2(O_p) \in U(\mathbb{F}_p)\}$$

for the upper triangular unipotent subgroup $U \subset G_1$. Then $[\Gamma, \Gamma'] = |G_1(\mathbb{F}_p)/U(\mathbb{F}_p)| = |\mathbf{P}^1(O/pO) \times (O/pO)^\times|$ which is prime to p. Since the action of Γ' on $L(k^*; \mathbb{F})$ factors through $U(\mathbb{F}_p)$, the evaluation $f(X, Y)$ at $(1, 0)$

gives a morphism of Γ'-modules $L(k^*;\mathbb{F}) \xrightarrow{\iota} \mathbb{F}$, where $\mathbb{F}$ is the constant Γ'-module with coefficients in $\mathbb{F}$. Similarly, $\mathbb{F} \xrightarrow{j} L(k^*;\mathbb{F})$ given by $\mathbb{F} \ni a \mapsto aY^{k^*}$ gives a morphism of Γ'-modules. We get two morphisms of cohomology groups:

$$H^g(\Gamma, L(k^*;\mathbb{F})) \xrightarrow{\iota} H^g(\Gamma,\mathbb{F}) \quad \text{and} \quad H^g(\Gamma,\mathbb{F}) \xrightarrow{\pi} H^g(\Gamma, L(k^*;\mathbb{F})),$$

which are induced by i and j, respectively. Choose $\tau \in M_2(F)$ which is sufficiently close to $\left(\begin{smallmatrix} 0 & -1 \\ p & 0 \end{smallmatrix}\right)$ p-adically and sufficiently close to $\left(\begin{smallmatrix} 1 & 0 \\ 0 & p \end{smallmatrix}\right)$ outside p so that τ normalizes Γ'. This is possible by the strong approximation theorem (e.g., [LFE] Theorem 3.2 for $F = \mathbb{Q}$ whose proof is valid even for $F \neq \mathbb{Q}$). By a simple computation, the endomorphism $j \circ i$ of the Γ'-module $L(k^*;\mathbb{F})$ is the action of τ. By the definition of the normalized Hecke operator $T(p)$ by the following commutative diagram,

$$\begin{array}{ccc} H^g(\Gamma', L(k^*;\mathbb{F})) & \xrightarrow{\tau} & H^g(\Gamma', L(k^*;\mathbb{F})) \\ {\scriptstyle \mathrm{Res}_{\Gamma/\Gamma'}}\uparrow & & \downarrow{\scriptstyle \mathrm{Tr}_{\Gamma/\Gamma'}} \\ H^g(\Gamma, L(k^*;\mathbb{F})) & \xrightarrow[T(p)]{} & H^g(\Gamma, L(k^*;\mathbb{F})) \end{array}$$

for the restriction $\mathrm{Res}_{\Gamma/\Gamma'}$ and the transfer map $\mathrm{Tr}_{\Gamma/\Gamma'}$ (cf. [MFG] 4.3.1), we find that $\mathrm{Tr}_{\Gamma/\Gamma'} \circ \pi \circ \iota \circ \mathrm{Res}_{\Gamma/\Gamma'} = T(p)$. This shows that $\iota \circ \mathrm{Res}_{\Gamma/\Gamma'}$ is injective over $H^g_{ord}(\Gamma, L(k^*;\mathbb{F}))$. Since $H^g(\Gamma',\mathbb{F})$ is finite-dimensional (because $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ is homologically equivalent to a finite simplicial complex; e.g., [H79] Section 3), the dimension of $H^g_{ord}(\Gamma', L(k^*;\mathbb{F}))$ is bounded by that of $H^g(\Gamma',\mathbb{F})$ independently of k. □

4.1.8 Moduli Problem of $\Gamma_1^1(\mathfrak{N})$-Type

Recall open compact subgroups $\widehat{\Gamma}_1(\mathfrak{N})$, $\widehat{\Gamma}_1^1(\mathfrak{N})$, and $\widehat{\Gamma}(\mathfrak{N})$ in $G(\widehat{\mathbb{Z}}) = GL_2(\widehat{O})$ defined in (4.5). As long as the moduli problem of type $\Gamma_1^1(\mathfrak{N})$ is representable by a quasi-projective smooth scheme $\mathfrak{M}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))_{/\mathbb{Z}_{(p)}}$, for its smooth projective toroidal compactification $M(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))_{/\mathbb{Z}_{(p)}}$, we get the same assertions as in Theorem 4.9 replacing the pair $(\mathfrak{M}(\mathfrak{c}, \Gamma(N)), M(\mathfrak{c}, \Gamma(N)))$ by $(\mathfrak{M}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N})), M(\mathfrak{c}, \Gamma_1^1(\mathfrak{N})))$ for weights k with $k \in (\mathbb{Z}\cdot I + 2\mathbb{Z}[I])$ (i.e., $k_\sigma \equiv k_\tau \bmod 2$ for all $\sigma, \tau : F \hookrightarrow \overline{\mathbb{Q}}$). This parity condition is necessary to define $\epsilon \mapsto \epsilon^{k/2}$ for totally positive units $\epsilon \in T(\mathbb{Z})$ (since $\epsilon^I = 1$ for such units). We only need the representability of the $\Gamma_1^1(\mathfrak{l}^m\mathfrak{N})$-moduli problem for a sufficiently large m and a prime $\mathfrak{l}$ to have a result similar to Theorem 4.9 as we describe this fact now.

To formulate the result, we consider the moduli problem of type Γ for $\Gamma = \Gamma_1^1(\mathfrak{N})$ and $\Gamma_1(\mathfrak{N})$, and accordingly, when representable, we define D_ℓ^{cusp}, V, $\mathcal{V}$, $\mathcal{V}_{cusp}^{ord,*}$, and $\mathcal{V}_{cusp}^{ord}$ with the level Γ-structure. Even if the modulus problem is not representable, we have the coarse moduli scheme $\mathfrak{M}(\mathfrak{c}, \Gamma)_{/\mathcal{W}}$ by making a geometric quotient (cf. [GME] Proposition 1.8.4) of $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ (over $\mathcal{W}$)

by a finite constant group $\widehat{\Gamma}/\widehat{\Gamma}(N)$ for a suitable $N \geq 3$ as in the proof of the following theorem. Then we define the above objects by the submodules (of the corresponding module of level $\Gamma(N)$) fixed by $\widehat{\Gamma}/\widehat{\Gamma}(N)$. When it is necessary to indicate their dependence on $(\mathfrak{c}, \Gamma)$, we write $D_\ell^{cusp}(\mathfrak{c}, \Gamma)$, $V(\mathfrak{c}, \Gamma)$, $\mathcal{V}(\mathfrak{c}, \Gamma)$, $\mathcal{V}_{cusp}^{ord,*}(\mathfrak{c}, \Gamma)$, and $\mathcal{V}_{cusp}^{ord}(\mathfrak{c}, \Gamma)$ in place of D_ℓ^{cusp}, V, $\mathcal{V}$, $\mathcal{V}_{cusp}^{ord,*}$, and $\mathcal{V}_{cusp}^{ord}$.

Theorem 4.10 *Let $\mathfrak{N}$ be an integral ideal prime to p with $\mathfrak{N} \cap \mathbb{Z} = (N)$. Suppose that p is prime to $Nd(F)\mathfrak{c}$ and that $p \geq 5$ if $N = 1$. Then we have for $T = \mathrm{Res}_{O/\mathbb{Z}}\mathbb{G}_m$,*

(1) *The submodule $D_\ell^{cusp}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))$ is dense in the p-adically complete space $V_{cusp}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))$;*
(2) *The dual $\mathcal{V}_{cusp}^{ord,*}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N})) = \mathrm{Hom}_W(V_{cusp}^{ord}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N})), W)$ of the module $\mathcal{V}_{cusp}^{ord}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))$ is a projective $W[[T(\mathbb{Z}_p)]]$-module of finite type;*
(3) *$\mathcal{V}_{cusp}^{ord,*}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N})) \otimes_{W[[T(\mathbb{Z}_p)]],k} W \cong \mathrm{Hom}_W(H_{ord}^0(M(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))[\frac{1}{E}], \underline{\omega}_k), W)$ if $k \geq 3I$;*
(4) *If $k \geq 3I$, the projector e induces an isomorphism*

$$H_{ord}^0(M(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))[\frac{1}{E}], \underline{\omega}_k) \cong H_{ord}^0(M(\mathfrak{c}, \Gamma_1^1(\mathfrak{N})), \underline{\omega}_k).$$

Proof. We deduce the theorem from Theorem 4.9. First suppose that $\mathcal{W}$ contains a primitive Nth root of unity ζ. For the differental idele $d \in F_{\mathbb{A}}^\times$ with $d^{(\mathfrak{d})} = 1$ and $d\widehat{O} = \widehat{\mathfrak{d}}$, we identify $\mathfrak{d}^{-1}/N\mathfrak{d}^{-1} = \mathfrak{d}^{-1} \otimes \mathbb{Z}/N\mathbb{Z}$ with O/NO by $x \mapsto dx$. By $m \mapsto \zeta^m$, we have an isomorphism $O/NO \cong \mathfrak{d}^{-1} \otimes \mu_N$ well-defined over $\mathcal{W}$. Then $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ over $\mathcal{W}$ is the moduli scheme classifying test objects $(A, \lambda, \phi : (O/NO)^2 \cong A[N])_{/R}$ with $\langle \phi(1,0), \phi(0,1) \rangle_N = \zeta$ for $\mathcal{W}$-algebras R. Thus we have a natural action of $\gamma \in SL_2(O/NO)$ on $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ induced by the functorial action on the test objects $(A, \lambda, \phi) \mapsto (A, \lambda, \phi \circ \gamma)$. The coarse moduli scheme $\mathfrak{M}(\mathfrak{c}, \Gamma_1^1(N))_{/\mathcal{W}}$ of $\Gamma_1^1(N)$-type is given by the geometric quotient $\mathfrak{M}(\mathfrak{c}, \Gamma(N))/U$ (cf. [GME] Proposition 1.8.4) by the constant subgroup of order prime to p, $U = \left\{ \left(\begin{smallmatrix} 1 & u \\ 0 & 1 \end{smallmatrix}\right) \middle| u \in O/NO \right\} \subset SL_2(O/NO)$. By the definition of the quotient, we have

$$X(\mathfrak{c}, \Gamma_1^1(N)) = H^0(U, X) \quad \text{for } X = D_\ell^{cusp},\ V,\ \mathcal{V},\ \mathcal{V}_{cusp}^{ord,*}, \text{ and } \mathcal{V}_{cusp}^{ord},$$

where $X(\mathfrak{c}, \Gamma_1^1(N))$ is the object corresponding to X defined relative to the $\Gamma_1^1(N)$-moduli problems. Note that $X(\mathfrak{c}, \Gamma_1^1(N))$ is a direct summand of X (with projection $\pi : X \to X(\mathfrak{c}, \Gamma_1^1(N))$ given by $\pi(x) = |U|^{-1} \sum_{u \in U} u(x)$), because $|U|$ is prime to p. Thus each assertion of the theorem follows from the corresponding assertion of Theorem 4.9 when $\mathfrak{N} = (N)$ with $N \geq 3$. For a general $\mathcal{W}$ that may not contain ζ, the object we studied over $\mathcal{W}[\zeta]$ is the scalar extension to $\mathcal{W}[\zeta]$ over $\mathcal{W}$; so, by faithfully flat descent, we recover the desired result for $\mathcal{W}$ when $\mathfrak{N} = (N)$.

We now treat the case where $\mathfrak{N}$ is not necessarily generated by an integer $N \geq 3$. The $\Gamma_1^1(\mathfrak{N})$-moduli problem is representable if each test object

of level $\Gamma_1^1(\mathfrak{N})$ does not have any automorphism other than the identity, because then the quotient process making $\mathfrak{M}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))$ out of $\mathfrak{M}(\mathfrak{c}, \Gamma(N))_{/\mathcal{W}}$ at the functor side and the moduli side matches. Suppose $\mathfrak{N} \subsetneq O$, and choose a prime factor $\mathfrak{l}$ of $\mathfrak{N}$. By the positivity of Rosati involution ([ABV] Section 21), the automorphism group of a test object (A, λ) is finite. By classification of the endomorphism algebra of a given abelian variety (cf. [ABV] Application I in Section 21), the size $|\operatorname{Aut}(A, \lambda)|$ is bounded independently of the test object. Thus for a sufficiently large power $\mathfrak{l}^m$, $\mathfrak{M}(\mathfrak{c}, \Gamma_1^1(\mathfrak{l}^m\mathfrak{N}))$ is representable over $\mathcal{W}$, because then each test object of level $\Gamma_1^1(\mathfrak{l}^m\mathfrak{N})$ does not have any automorphism other than the identity map. Thus the theorem is valid for level $\Gamma_1^1(\mathfrak{l}^m\mathfrak{N})$. Note that $\mathfrak{M}(\mathfrak{c}, \Gamma_1^1(\mathfrak{l}^m\mathfrak{N}))$ is a faithfully flat finite covering of $\mathfrak{M}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))$ of degree d prime to p (actually d is a power of $N(\mathfrak{l})$). Since the restriction composed with the trace map is the multiplication by the degree d, we can realize all objects $D_\ell^{cusp}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))$, $V(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))$, $\mathcal{V}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))$, $\mathcal{V}_{cusp}^{ord,*}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))$, and $\mathcal{V}_{cusp}^{ord}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))$ as a $W[[T(\mathbb{Z}_p)]]$-direct summand of the corresponding objects of level $\Gamma_1^1(\mathfrak{l}^m\mathfrak{N})$. By this, we conclude the desired assertions if $\mathfrak{N} \subsetneq O$.

We now suppose that $\mathfrak{N} = O$. Since p is unramified in $F/\mathbb{Q}$ and $p \geq 5$, $\operatorname{Gal}(F[\mu_p]/F) \cong (\mathbb{Z}/p\mathbb{Z})^\times$ by sending the Frobenius element at a prime $\mathfrak{l} \nmid p$ to $N(\mathfrak{l}) \mod p$ (class field theory). Thus we find a prime $\mathfrak{l}$ of F such that $N(\mathfrak{l}) \mod p$ generates the cyclic group $(\mathbb{Z}/p\mathbb{Z})^\times$. In particular $p \nmid (N(\mathfrak{l}) \pm 1)$. Then the modulus variety $\mathfrak{M}(\mathfrak{c}, \Gamma(\mathfrak{l}))$ for $\widehat{\Gamma}$ is a faithfully flat finite covering of degree prime to p over $\mathfrak{M}(\mathfrak{c}, \Gamma(1))$ (the degree is a factor of $N(\mathfrak{l})(N(\mathfrak{l})-1)(N(\mathfrak{l})^2-1)$). Then by the same argument as above, we get the result for $\mathfrak{N} = O$. □

4.1.9 p-Adic Modular Forms on $PGL(2)$

We now interpret the result obtained in the previous section in terms of automorphic forms on $PG = \operatorname{Res}_{O/\mathbb{Z}} PGL(2)$. For simplicity, we hereafter assume that k is even, since the general case has been explained in [SGL] Chapter 2.

We consider a test object $(A, O_+^\times\lambda, i)_{/R}$ of level $\Gamma_1(\mathfrak{N})$. For any ideal $\mathfrak{z}$ prime to Np ($(N) = \mathfrak{N} \cap \mathbb{Z}$), we make a quotient $A' = A/A[\mathfrak{z}] = A \otimes_O \mathfrak{z}^{-1}$; thus, $A = A' \otimes \mathfrak{z}$. Then $A'[\mathfrak{N}] \cong A[\mathfrak{N}]$ canonically; so, i induces a level $\Gamma_1(\mathfrak{N})$-structure i' into A'. Let $P(A) = \{\lambda \in \operatorname{Hom}(A, {}^tA) | {}^t\lambda = \lambda\}$ and $P_+(A) \subset P(A)$ be the subset made of polarizations. Then we have an isomorphism $\lambda : P \cong \mathfrak{c}$ of O-modules taking P_+ onto the subset $\mathfrak{c}_+$ of totally positive elements of $\mathfrak{c}$. Dualizing the exact sequence $0 \to A[\mathfrak{z}] \to A \to A' \to 0$, we get another exact sequence $0 \to {}^tA'[\mathfrak{z}] \to {}^tA' \to {}^tA \to 0$, because ${}^tA'[\mathfrak{z}]$ is the Cartier dual of $A[\mathfrak{z}]$ (see the proof of Theorem 8.9). This shows ${}^tA' \cong {}^tA \otimes \mathfrak{z} \cong A \otimes \mathfrak{c}\mathfrak{z} \cong A' \otimes \mathfrak{c}\mathfrak{z}^2$; so, λ induces $\lambda' : P(A')_+ \cong (\mathfrak{c}\mathfrak{z}^2)_+$. Thus $(A, O_+^\times\lambda, i) \mapsto (A', O_+^\times\lambda', i')$ induces $\langle\mathfrak{z}\rangle : V(\mathfrak{c}\mathfrak{z}^2, \Gamma_1(\mathfrak{N})) \cong V(\mathfrak{c}, \Gamma_1(\mathfrak{N}))$ by $f|\langle\mathfrak{z}\rangle(A, O_+^\times\lambda, i) = f(A', O_+^\times\lambda', i')$. The operator $\langle\mathfrak{z}\rangle$ is called the *diamond* operator. We identify $V(\mathfrak{c}, \Gamma_1(\mathfrak{N}))$ and $V(\mathfrak{c}\mathfrak{z}^2, \Gamma_1(\mathfrak{N}))$ (resp. $\mathcal{V}(\mathfrak{c}, \Gamma_1(\mathfrak{N}))$ and $\mathcal{V}(\mathfrak{c}\mathfrak{z}^2, \Gamma_1(\mathfrak{N}))$) by $\langle\mathfrak{z}\rangle$, and write it as $V_\mathfrak{c}$ (resp. $\mathcal{V}_\mathfrak{c}$) whose ordinary and cuspidal part is written as $V_{cusp,\mathfrak{c}}^{ord}$ (resp.

$\mathcal{V}^{ord}_{cusp,\mathfrak{c}}$). Thus these spaces only depend on the strict ideal class of $\mathfrak{c}$ modulo square ideal classes. We then define

$$\mathcal{V}^{ord}_{cusp}(\widehat{\Gamma}_1(\mathfrak{N})) = \bigoplus_{\mathfrak{c}} \mathcal{V}^{ord}_{cusp,\mathfrak{c}} \quad \text{and} \quad \mathcal{V}^{ord,*}_{cusp}(\widehat{\Gamma}_1(\mathfrak{N})) = \bigoplus_{\mathfrak{c}} \mathcal{V}^{ord,*}_{cusp,\mathfrak{c}}, \tag{4.15}$$

where $\mathfrak{c}$ runs over strict equivalence classes of ideals modulo square classes; thus, it runs over the group $Cl^+_F/(Cl^+_F)^2$, where Cl^+_F is the strict ideal class group. Recall that

$$PGL_2(F_{\mathbb{A}}) = \bigsqcup_{c \in Cl^+_F/(Cl^+_F)^2} PGL_2(F) \left(\begin{smallmatrix} c & 0 \\ 0 & 1 \end{smallmatrix}\right) \overline{\Gamma}_1(\mathfrak{N}) PGL_2^+(F_\infty),$$

where c runs over a complete representative set for $Cl^+_F/(Cl^+_F)^2$ in $F^\times_{\mathbb{A}}$; $F_\infty = F \otimes_{\mathbb{Q}} \mathbb{R}$; $PGL_2^+(F_\infty)$ is the identity-connected component of $PGL_2(F_\infty)$, and $\overline{\Gamma}_1(\mathfrak{N})$ is the image of $\widehat{\Gamma}_1(\mathfrak{N})$ in $PGL_2(F^{(\infty)}_{\mathbb{A}})$. Thus we may regard $\mathcal{V}^{ord,*}_{cusp}(\widehat{\Gamma}_1(\mathfrak{N}))$ as the W-dual of the space of p-adic cusp forms of level $\widehat{\Gamma}_1(\mathfrak{N})$ on $PGL_2(F_{\mathbb{A}})$. We now repeat the vertical control theorem in the setting of $PGL(2)$.

Theorem 4.11 *Let $\mathfrak{N}$ be an integral ideal prime to p with $\mathfrak{N} \cap \mathbb{Z} = (N)$. Suppose that p is prime to $Nd(F)\mathfrak{c}$ and $p \geq 5$. Then we have for $T = \mathrm{Res}_{O/\mathbb{Z}}\mathbb{G}_m$,*

(1) *The dual $\mathcal{V}^{ord,*}_{cusp}(\widehat{\Gamma}_1(\mathfrak{N}))$ of $\mathcal{V}^{ord}_{cusp}(\widehat{\Gamma}_1(\mathfrak{N}))$ is a projective $W[[T(\mathbb{Z}_p)]]$-module of finite type;*
(2) *If $k \geq 3I$, $\mathcal{V}^{ord,*}_{cusp}(\widehat{\Gamma}_1(\mathfrak{N})) \otimes_{W[[T(\mathbb{Z}_p)]],k} W$ is isomorphic to the direct sum $\bigoplus_{\mathfrak{c} \in Cl^+_F/(Cl^+_F)^2} \mathrm{Hom}_W(H^0_{ord}(M(\mathfrak{c}, \Gamma_1(\mathfrak{N})), \underline{\omega}_k), W)$,*

where $M(\mathfrak{c}, \Gamma_1(\mathfrak{N}))$ is the smooth projective toroidal compactification of the moduli scheme $\mathfrak{M}(\mathfrak{c}, \Gamma_1(\mathfrak{N}))$.

Proof. Choose a prime $\mathfrak{l} \nmid pN$ with $p \nmid (N(\mathfrak{l})^2 - 1)$. We consider the mixed moduli problem $\Gamma^1_1(\mathfrak{l}^m) \cap \Gamma_1(\mathfrak{N})$, whose test objects are $(A, \lambda, \phi_{\mathfrak{l}}, i_{\mathfrak{N}})_{/R}$ with $\mathfrak{c}$-polarization λ, a level $\Gamma_1(\mathfrak{N})$-structure $i_{\mathfrak{N}}$ and a level $\Gamma^1_1(\mathfrak{l}^m)$-structure $\phi_{\mathfrak{l}}$. This problem is representable by a moduli scheme $\mathfrak{M} = \mathfrak{M}(\mathfrak{c}, \Gamma^1_1(\mathfrak{l}^m) \cap \Gamma_1(\mathfrak{N}))$ for sufficiently large m. Then $\mathfrak{M}$ is a faithfully flat finite covering of the coarse modulus $\mathfrak{M}(\mathfrak{c}, \Gamma_1(\mathfrak{N}))$ of degree prime to p. Since the result for each piece $\mathcal{V}^{ord,*}_{cusp,\mathfrak{c}}$ is valid for $\mathfrak{M}$, the result can be proven as in the proof of Theorem 4.10. □

A subgroup C of an abelian scheme $A_{/S}$ is called *étale cyclic* of order $\mathfrak{N}$ if C is isomorphic to the constant group $(O/\mathfrak{N})_{/S}$ étale locally. In place of level $\Gamma_1(\mathfrak{N})$-structure, we can think of a quadruple

$$(A, \lambda, i_p : \mu_{p^\infty} \otimes \mathfrak{d}^{-1} \hookrightarrow A[p^\infty], C)$$

for an étale cyclic subgroup C of order $\mathfrak{N}$. The level structure C is called the level $\Gamma_0(\mathfrak{N})$-structure. Considering this type of level structure, we may construct $\mathcal{V}_{cusp}(\widehat{\Gamma}_0(\mathfrak{N}))$ and $V_{cusp}(\widehat{\Gamma}_0(\mathfrak{N}))$ in the same manner as above, replacing level $\Gamma_1(\mathfrak{N})$-structures by those of level $\Gamma_0(\mathfrak{N})$.

More generally, identifying $\mu_N \otimes O^*$ with $O/(N)$ (for N prime to p), we can let $g \in GL_2(O/(N))$ act on level $\Gamma(N)$-structure $\phi_N : (O/(N))^2 \cong A[N]$ by $\phi_N \mapsto \phi_N \circ g$. Thus for a given subgroup $\Gamma \subset GL_2(\widehat{O})$, we can think of the orbit $\overline{\phi} = \phi_N \circ \Gamma$ and can think of the moduli problem classifying $(A, \lambda, \overline{\phi})$. We call this level structure $\overline{\phi}$ of type Γ. We write $\mathfrak{M}(\mathfrak{c}, \Gamma)$ for the coarse moduli scheme for this modulus problem of type Γ with the $\mathfrak{c}$-polarization requirement. We again define the module $\mathcal{V}^{ord}_{cusp}(\mathfrak{c}, \Gamma)$ of p-adic modular forms for this modulus problem as before.

Since the stabilizer of the standard subgroup $(O/\mathfrak{N} \oplus 0)$ of $(O/\mathfrak{N})^2$ is given by the subgroup $\widehat{\Gamma}_0(\mathfrak{N})$ of $GL_2(\widehat{O})$, the level $\Gamma_0(\mathfrak{N})$-structure introduced above can be considered to be the level structure of type $\Gamma_0(\mathfrak{N})$. Any level $\Gamma_0(\mathfrak{N})$-structure (even if we require isomorphisms to keep polarization) has automorphisms ± 1; so, it cannot be representable without adding an extra structure. Nevertheless, we obtain the following theorem in exactly the same manner as in the proof of Theorem 4.11.

Theorem 4.12 *Let the notation be as above. Suppose that $\mathfrak{N}$ is an integral ideal prime to p with $\mathfrak{N} \cap \mathbb{Z} = (N)$ and that $p \geq 5$ is prime to $Nd(F)\mathfrak{c}$. Then we have for $T = \mathrm{Res}_{O/\mathbb{Z}}\mathbb{G}_m$,*

(1) *The dual $\mathcal{V}^{ord,*}_{cusp}(\widehat{\Gamma}_0(\mathfrak{N}))$ of $\mathcal{V}^{ord}_{cusp}(\widehat{\Gamma}_0(\mathfrak{N}))$ is a projective $W[[T(\mathbb{Z}_p)]]$-module of finite type;*
(2) *If $k \geq 3I$, $\mathcal{V}^{ord,*}_{cusp}(\widehat{\Gamma}_0(\mathfrak{N})) \otimes_{W[[T(\mathbb{Z}_p)]],k} W$ is isomorphic to the direct sum $\bigoplus_{\mathfrak{c} \in Cl^+_F/(Cl^+_F)^2} \mathrm{Hom}_W(H^0_{ord}(M(\mathfrak{c}, \Gamma_0(\mathfrak{N})), \underline{\omega}_k), W)$, where $M(\mathfrak{c}, \Gamma_0(\mathfrak{N}))$ is the smooth toroidal compactification of $\mathfrak{M}(\mathfrak{c}, \Gamma_0(\mathfrak{N}))$.*

4.1.10 Hecke Operators on Geometric Modular Forms

We call a ring R a *p-adic ring* if $R = \varprojlim_n R/p^n R$. We now assume that the base algebra R is a $\mathcal{W}$-algebra when we consider classical modular forms. When we consider p-adic modular forms, we further suppose that R is a p-adic ring; so, it is a W-algebra automatically for $W = \varprojlim_n \mathcal{W}/p^n\mathcal{W}$.

For a prime ideal $\mathfrak{q}$ of F, we define Hecke operators $U(\mathfrak{q}^r)$ and $T(1, \mathfrak{q}^r)$. The operator $T(1, \mathfrak{q})$ is often written as $T(\mathfrak{q})$, but $T(1, \mathfrak{q}^r) \neq T(\mathfrak{q}^r)$ if $r > 1$. If $\mathfrak{l}$ and $\mathfrak{q}$ are distinct prime ideals, we see easily from the definition that the operators defined for $\mathfrak{l}$ and $\mathfrak{q}$ commute. The operator we need (to define the projectors e and e°) is $U(p) = \prod_{\mathfrak{p}|p} U(\mathfrak{p})$ and $T(p) = \prod_{\mathfrak{p}|p} T(\mathfrak{p})$, respectively.

We call a subgroup C of an abelian scheme $A_{/S}$ *cyclic* of order $\mathfrak{q}^r$ if either $C \cong O/\mathfrak{q}^r$ or $C \cong \mu_{\mathfrak{q}^r}$ over an étale faithfully flat extension of S. We call C *étale cyclic* if $C \cong O/\mathfrak{q}^r$ étale locally. We think of test objects $(A, \lambda, i, \omega)_{/S}$ for $(A, \lambda, i)_{/S}$ giving a point of $\mathfrak{M}(\mathfrak{c}, \Gamma^1_1(\mathfrak{N}))(S)$, where ω is a generator of $H^0(A, \Omega_{A/S})$ over $\mathcal{O}_S \otimes_{\mathbb{Z}} O$. Assume that S is an R-scheme for a base ring R. When $\mathfrak{q}|p\mathfrak{N}$, we have additional information C of a cyclic subgroup scheme $C \subset A$ of order $\mathfrak{q}$ given by $i(\mu_\mathfrak{q})$ if $\mathfrak{q} \nmid p$ and the unique connected subgroup $C = i(\mu_\mathfrak{q})$ of $A[\mathfrak{q}]$ if $\mathfrak{q}|p$ as long as the base ring R is a p-adic ring. Here

"connectedness" is defined as a scheme relative to $Spec(R)$; so, a connected scheme over $Spec(R)$ has a connected topological space over each connected component of $Spec(R)$. Since $A_{/R}$ is an O-module, we may consider $A[\mathfrak{a}] = \{x \in A | \mathfrak{a}x = 0\}$ for an O-ideal $\mathfrak{a} \neq 0$. For the integer a with $\mathfrak{a} \cap \mathbb{Z} = (a)$, $A[\mathfrak{a}] \subset A[a]$. The scheme $A[a] = \mathrm{Ker}(a : A \to A)$ is a locally free group scheme of finite rank, because A is smooth and proper over R (see [GME] 2.8.1). Choosing a generator α of the ideal $\mathfrak{a}/aO$, the group scheme $A[\mathfrak{a}]$ is given by the kernel of α in $A[a]$, and hence $A[\mathfrak{a}]$ is locally free of finite rank.

We start with describing the definition of Hecke operators for classical modular forms, and later we adjust the definition to p-adic modular forms. We define Hecke operators $T(1, \mathfrak{q}^r)$ (if $\mathfrak{q} \nmid \mathfrak{N}$) and $U(\mathfrak{q}^r)$ (if $\mathfrak{q}|\mathfrak{N}$) on the space of modular forms $G_k(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}); R)$ first under the assumption that $N(\mathfrak{q})$ is invertible in the base ring R. For an étale cyclic subgroup C' of $A_{/S}$ of order $\mathfrak{q}^r$, we can make the quotient abelian scheme $A' = A/C'$ with the projection $\pi : A \to A'$ (cf. [ABV] Section 12 and [GME] Proposition 1.8.4). The $\mathfrak{c}$-polarization λ induces a $\mathfrak{c}\mathfrak{q}^r$-polarization $\pi_*\lambda$. This can be checked as follows. Tensoring the exact sequence $\mathfrak{q}^r \hookrightarrow O \twoheadrightarrow O/\mathfrak{q}^r$ with ${}^tA \overset{\lambda}{=} A \otimes \mathfrak{c}$, we have another exact sequence $A \otimes \mathfrak{c}\mathfrak{q}^r[\mathfrak{q}^r] \hookrightarrow A \otimes \mathfrak{c}\mathfrak{q}^r \twoheadrightarrow A \otimes \mathfrak{c}$. Taking the dual of $\pi : A \to A'$, we have one more exact sequence $\widehat{C}' \hookrightarrow {}^tA' \twoheadrightarrow A$ for the Cartier dual $\widehat{C}'$ of C', which gives rise to a short exact sequence $\widehat{C}' \hookrightarrow {}^tA'[\mathfrak{q}^r] \xrightarrow{\lambda \circ {}^t\pi} C' \otimes \mathfrak{c}$. The kernel of the composite $(\pi \otimes \mathrm{id}) \circ \lambda \circ {}^t\pi : {}^tA' \to A \otimes \mathfrak{c}$ is the entire $\mathfrak{q}^r$-torsion subgroup ${}^tA'[\mathfrak{q}^r]$. Since ${}^tA'/{}^tA'[\mathfrak{q}^r] = {}^tA' \otimes \mathfrak{q}^{-r}$, we have constructed an isomorphism $(\pi \otimes \mathrm{id}) \circ \lambda \circ {}^t\pi : {}^tA' \otimes \mathfrak{q}^{-r} \cong A \otimes \mathfrak{c}$. Tensoring $\mathfrak{q}^r$ with this isomorphism, we get the desired $\lambda' : {}^tA' \cong A' \otimes \mathfrak{c}\mathfrak{q}^r$.

Since the projection $\pi : A \to A/C'$ is étale ($\Leftrightarrow \pi^* : \Omega_{(A/C')/S} \cong \Omega_{A/S}$), the differential ω gives a differential $(\pi^*)^{-1}\omega$ on A/C'. If $\mathfrak{q}|\mathfrak{N}$ and the schematic intersection $C' \cap C = C' \times_A C$ is reduced to the identity $\{0\}$ (in this case, we say that C' and C are *disjoint*), $\pi \circ i$ gives rise to the level $\Gamma_1^1(\mathfrak{N})$-structure on A/C'. Then we define for $f \in G_k(\mathfrak{c}\mathfrak{q}^r, \Gamma_1^1(\mathfrak{N}); R)$,

$$f|U(\mathfrak{q}^r)(A, \lambda, i, \omega) = \frac{1}{N(\mathfrak{q}^r)} \sum_{C'} f(A/C', \pi_*\lambda, \pi \circ i, (\pi^*)^{-1}\omega), \tag{4.16}$$

where C' runs over all étale cyclic subgroups of order $\mathfrak{q}^r$ disjoint from C.

Since $\pi_*\lambda = \pi \circ \lambda \circ {}^t\pi$ is a $\mathfrak{c}\mathfrak{q}^r$-polarization, the modular form f has to be defined for abelian varieties with $\mathfrak{c}\mathfrak{q}^r$-polarization. When $\mathfrak{q} \nmid \mathfrak{N}$ and $N(\mathfrak{q})$ is invertible in R, we define the action of the *Hecke operator* $T(1, \mathfrak{q}^r)$ on $f \in G_k(\mathfrak{c}\mathfrak{q}^r, \Gamma_1^1(\mathfrak{N}); R)$ by

$$f|T(1, \mathfrak{q}^r)(A, \lambda, i, \omega) = \frac{1}{N(\mathfrak{q}^r)} \sum_{C'} f(A/C', \pi_*\lambda, \pi \circ i, (\pi^*)^{-1}\omega), \tag{4.17}$$

where C' runs over all étale cyclic subgroups of order $\mathfrak{q}^r$. We can check that $f|U(\mathfrak{q}^r)$ and $f|T(1, \mathfrak{q}^r)$ stay in $G_k(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}); R)$ if $f \in G_k(\mathfrak{c}\mathfrak{q}^r, \Gamma_1^1(\mathfrak{N}); R)$.

Now we allow the ring R in which $N(\mathfrak{q})$ may not be invertible. First assume that $R\left[\frac{1}{N(\mathfrak{q})}\right]$ is flat over R. Then by the flat base change theorem ([ALG] Proposition III.9.3), we have

$$G_k(\mathfrak{c},\Gamma;R\left[\frac{1}{N(\mathfrak{q})}\right]) = H^0(\mathfrak{M}(\mathfrak{c},\Gamma),\underline{\omega}^{kI}_{/R[\frac{1}{N(\mathfrak{q})}]})$$
$$\cong H^0(\mathfrak{M}(\mathfrak{c},\Gamma),\underline{\omega}^{kI}_{/R})\otimes_R R\left[\frac{1}{N(\mathfrak{q})}\right] = G_k(\mathfrak{c},\Gamma;R)\otimes_R R\left[\frac{1}{N(\mathfrak{q})}\right].$$

The operators are well-defined over $R\left[\frac{1}{N(\mathfrak{q})}\right]$. Thus if it preserves the R-integral structure, it is well-defined over R. The operator $U(\mathfrak{q}^r)$ is always integral and $T(1,\mathfrak{q}^r)$ is integral if the weight k is sufficiently positive. We show in Section 8.3.1 the R-integrality of these operators using the Serre–Tate deformation theory when R is a p-adic W-algebra. The integrality can be also checked using the q-expansion principle (which shows that the sum in (4.16) is actually divisible by $N(\mathfrak{q}^r)$ and $T(1,\mathfrak{p}) \equiv U(\mathfrak{p}) \mod p$ as long as $k \geq 2I$; see Section 4.2.9).

If $\underline{\omega}^{kI}$ is very ample (which holds true if $k \gg 0$), the sheaf is generated by global sections; so, $H^0(\mathfrak{M},\underline{\omega}^{kI}_{R/p^mR}) = H^0(\mathfrak{M},\underline{\omega}^{kI}_{/R})\otimes R/p^mR$, and

$$G_k(\mathfrak{c},\Gamma;R/p^mR) = H^0\left(\mathfrak{M}(\mathfrak{c},\Gamma),\underline{\omega}^{kI}_{/R/p^mR}\right)$$
$$\cong H^0\left(\mathfrak{M}(\mathfrak{c},\Gamma),\underline{\omega}^{kI}_{/R}\right)\otimes_R R/p^mR = G_k(\mathfrak{c},\Gamma;R)\otimes_R R/p^mR$$

follows from the long exact sequence from the short one $R \xrightarrow{p^m} R \twoheadrightarrow R/p^mR$. Thus if the operators $U(\mathfrak{q}^r)$ and $T(1,\mathfrak{q}^r)$ are R-integral, they are well-defined over R/p^mR. Later we give a definition of these operators as an algebraic correspondence, which works well over a more general ring R (see Section 4.2.5).

We now deal with p-adic modular forms defined over a p-adic ring R classifying test objects $(A,\lambda,i_p:\mu_{p^\infty}\otimes\mathfrak{d}^{-1}\hookrightarrow A[p^\infty],\phi)_{/R'}$ (for R-algebras R') with prime-to-p level structure ϕ of type Γ. When $\mathfrak{q}$ is prime to p, $N(\mathfrak{q})$ is invertible in a p-adic ring; so, we define $f|U(\mathfrak{q}^r)$ (for $\mathfrak{q}|\mathfrak{N}$) and $f|T(1,\mathfrak{q}^r)$ (for $\mathfrak{q}\nmid\mathfrak{N}$) replacing the differential $(\pi^*)^{-1}\omega$ by the level p^∞-structure $\pi\circ i_p$ in the above definitions (4.16) and (4.17).

To define $U(\mathfrak{p})$ $(\mathfrak{p}|p)$ acting on p-adic modular forms, let $S_{\infty/W}$ be the formal completion of $\mathfrak{M}(\mathfrak{c},\Gamma)[\frac{1}{E}]$ along its modulo p fiber for the lift of the Hasse invariant as in (E). Let $(\mathbb{X},\boldsymbol{\lambda},\boldsymbol{\phi})_{/\mathfrak{M}(\mathfrak{c},\Gamma)}$ be the universal test object (supposing the classification functor of prime-to-p level Γ is representable). We pick an étale cyclic subgroup $\mathcal{C}'\subset\mathbb{X}[\mathfrak{p}]_{/S_\infty}$ of order $\mathfrak{p}$. As we show in Section 8.3.1 (particularly, the argument below (8.25)), $\mathcal{C}'$ can be defined only over a locally free covering S'_∞/S_∞ of rank $N(\mathfrak{p})$, and S'_∞/S_∞ is radiciel (purely inseparable over the generic point of S_∞). By the universality, we have a unique morphism $Spec(R)\xrightarrow{\varphi}S_\infty\subset\mathfrak{M}(\mathfrak{c},\Gamma)$ with $\varphi^*(\mathbb{X},\boldsymbol{\lambda},\boldsymbol{\phi})\cong(A,\lambda,\phi)$.

Then $Spec(R') = \varphi^* S'_\infty$, and R' is an R-algebra (locally free of rank $N(\mathfrak{p})$) with trace map $\mathrm{Tr} : R' \to R$ (see [K] 3.11), and $C' = \varphi^* \mathcal{C}'$ gives an étale cyclic subgroup of $A_{/R'}$ of order $\mathfrak{p}$. Here the trace $\mathrm{Tr}(x)$ is given by the coefficient (times -1) of the characteristic polynomial over R of multiplication by x on R'. Since A/C' is defined over R', the operator $N(\mathfrak{p})U(\mathfrak{p})$ for a prime ideal $\mathfrak{p}|p$ can be defined by

$$f|N(\mathfrak{p})U(\mathfrak{p})(A, \lambda, i_p, \phi) = \mathrm{Tr}(f(A/C', \pi_*\lambda, \pi \circ i_p, \pi \circ \phi)). \tag{4.18}$$

If $R[\frac{1}{p}]$ is flat over R, the trace $\mathrm{Tr}(f(A/C', \pi_*\lambda, \pi \circ i_p, \pi \circ \phi)) \in R$ is just the sum $\sum_{C'} f(A/C', \pi_*\lambda, \pi \circ i_p, \pi \circ \phi)$ over all cyclic subgroups C' of order $\mathfrak{p}$ generically different from the connected component $A[\mathfrak{p}]^\circ_{/R}$ of $A[\mathfrak{p}]_{/R}$; so, (4.18) is compatible with (4.16). The divisibility of the operator $N(\mathfrak{p})U(\mathfrak{p})$ by $N(\mathfrak{p})$ can be checked by computing either the q-expansion of $f|N(\mathfrak{p})U(\mathfrak{p})$ (in Section 4.2.9) or its expansion at a Serre–Tate deformation space (as we do later in 8.3.1). The representability of the level Γ modulus problem is not essential in defining $U(\mathfrak{p})$, because we can define it at any level Γ if we once have the operator well-defined for a representable $\Gamma \cap \Gamma_1^1(\mathfrak{l}^m)$ moduli problem.

There is a shortcut to get the operator $U(\mathfrak{p})$ acting on p-adic modular forms if $R = W$. By the earlier argument for classical modular forms, we have a well-defined W-integral operator $U(\mathfrak{p})$ on $D_\ell^{cusp} \subset V^{cusp}$ in (4.13) (for any weight $\ell > 0$) given by (4.16). Then by the p-adic density of D_ℓ^{cusp} in V^{cusp} (Theorems 4.9 and 4.10), the operator $U(\mathfrak{p})$ extends to p-adic modular forms, which of course coincides with the one given by (4.18).

We have from the definition $U(\mathfrak{q}^r) = U(\mathfrak{q})^r$. If $\mathfrak{n} = (\xi)$ for a $\xi \in F_+^\times$, then $\lambda \mapsto \xi^{-1}\lambda$ gives a bijection between $\mathfrak{c}$-polarizations and $\mathfrak{n}\mathfrak{c}$-polarizations; so, it induces an identification $G_k(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}); R) \cong G_k(\mathfrak{n}\mathfrak{c}, \Gamma_1^1(\mathfrak{N}); R)$ by the association $(A, \lambda) \mapsto (A, \xi^{-1}\lambda)$. Thus we may regard $U(\xi) = \prod_{\mathfrak{q}} U(\mathfrak{q}^{e(\mathfrak{q})})$ for the prime decomposition $\mathfrak{n} = \prod_{\mathfrak{q}} \mathfrak{q}^{e(\mathfrak{q})}$ as an endomorphism of $G_k(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}); R)$. This really depends on the choice of ξ (because the identification is given by $\lambda \mapsto \xi^{-1}\lambda$), though $U(\xi)$ is well-defined, up to a unit scalar multiple, independently of the choice of ξ. Similarly, we can define $T(1, \xi)$ by $T(1, \xi) = \prod_{\mathfrak{q}} T(1, \mathfrak{q}^{e(\mathfrak{q})})$. Thus we have a well-defined $U(p)$ as an endomorphism of $G_k(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}); R)$. Similarly, we have $T(p)$ acting on $G_k(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}); R)$.

In the above definition, we can replace $\Gamma_1^1(\mathfrak{N})$ by $\Gamma_1(\mathfrak{N})$. Since the polarization of a level $\Gamma_1(\mathfrak{N})$ test object is specified only modulo multiplication by $O_+^\times$, $U(\xi)$ and $T(1, \xi)$ are well-defined independently of the generator ξ of $\mathfrak{n}$; thus, we may write $U(\mathfrak{n})$ and $T(1, \mathfrak{n})$ for $U(\xi)$ and $T(1, \xi)$, respectively.

Note that the space of modular forms $G_k(\mathfrak{c}, \mathfrak{N}; R)$ depends only on the strict ideal class of $\mathfrak{c}$. So we define

$$G_k(\widehat{\Gamma}_1(\mathfrak{N}); R) = \bigoplus_{\mathfrak{c} \in Cl_F^+} G_k(\mathfrak{c}, \mathfrak{N}; R), \quad V(\widehat{\Gamma}_1(\mathfrak{N}); R) = \bigoplus_{\mathfrak{c} \in Cl_F^+} V(\mathfrak{c}, \mathfrak{N}; R) \tag{4.19}$$

for the space of p-adic modular forms $V(\mathfrak{c}, \mathfrak{N}; R)$ over R. The Hecke operators $U(\mathfrak{q})$ and $T(1, \mathfrak{q}^r)$ permute the components $G_k(\mathfrak{c}, \mathfrak{N}; R)$ and give

endomorphisms of $G_k(\widehat{\Gamma}_1(\mathfrak{N}); R)$ (or $V(\widehat{\Gamma}_1(\mathfrak{N}); R)$). The diamond operator $\langle \mathfrak{z} \rangle : G_k(\mathfrak{c}, \mathfrak{N}; R) \to G_k(\mathfrak{c}\mathfrak{z}^2, \mathfrak{N}; R)$ (resp. $V(\mathfrak{c}, \mathfrak{N}; R) \to V(\mathfrak{c}\mathfrak{z}^2, \mathfrak{N}; R)$) introduced in the previous section for ideals $\mathfrak{z}$ prime to $p\mathfrak{N}$ also acts on the above spaces (permuting ideal classes $\mathfrak{c} \mapsto \mathfrak{c}\mathfrak{z}^2$). This is the easiest way of extending the definition of Hecke operators to $GL(2)$ and is based on the first identity of (4.7). There is another way of extending the definition of Hecke operators to $GL(2)$, using the open compact subgroup $S_1^1(\mathfrak{N})$ in (4.6) in place of $S_1(\mathfrak{N})$. Then the approximation theorem yields

$$\bigsqcup_{\mathfrak{c} \in Cl_F^+(\mathfrak{N})} \mathfrak{M}(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))(\mathbb{C}) = G(\mathbb{Q}) \backslash G(\mathbb{A}) / S_1^1(\mathfrak{N}) Z(\mathbb{R}) C_{\mathbf{0}}.$$

The above naive definition of the Hecke operator acting on $G_k(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}); R)$ gives rise to the Hecke operator on

$$G_k(\widehat{\Gamma}_1^1(\mathfrak{N}); R) = \bigoplus_{\mathfrak{c} \in Cl_F^+(\mathfrak{N})} G_k(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}); R)$$

$$\text{and} \quad V(\widehat{\Gamma}_1^1(\mathfrak{N}); R) = \bigoplus_{\mathfrak{c} \in Cl_F^+(\mathfrak{N})} V(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}); R)$$

as an endomorphism. Again $A \mapsto A \otimes \mathfrak{z}$ induces an operator $\langle \mathfrak{z} \rangle$ acting on the above spaces as long as $\mathfrak{z}$ is prime to $p\mathfrak{N}$.

The natural diagonal embeddings

$$G_k(\mathfrak{c}, \mathfrak{N}; R) \hookrightarrow \bigoplus_{\mathfrak{c}' \sim \mathfrak{c}} G_k(\mathfrak{c}', \Gamma_1^1(\mathfrak{N}); R) \quad \text{and} \quad V(\mathfrak{c}, \mathfrak{N}; R) \hookrightarrow \bigoplus_{\mathfrak{c}' \sim \mathfrak{c}} V(\mathfrak{c}', \Gamma_1^1(\mathfrak{N}); R)$$

are equivariant under the Hecke operator action. Here in the above sum, $\mathfrak{c}'$ runs over the classes in $Cl_F^+(\mathfrak{N})$ equivalent to $\mathfrak{c}$ in Cl_F^+.

Since $Cl_F^+(\mathfrak{N}) = (F_{\mathbb{A}}^{(\infty)})^\times / F_+^\times \det(\widehat{\Gamma}_1^1(\mathfrak{N}))$, we may define a quotient Cl_Γ of $Cl_F^+(\mathfrak{N})$ by $\det(\Gamma)$ for the more general subgroup $\Gamma \subset GL_2(\widehat{O})$ containing $\widehat{\Gamma}_1^1(\mathfrak{N})$. Then we define for a p-adic W-algebra R,

$$V(\Gamma; R) = \bigoplus_{\mathfrak{c} \in Cl_\Gamma} V(\mathfrak{c}, \Gamma; R), \tag{4.20}$$

on which Hecke operators act as we show later.

4.2 Hilbert Modular Shimura Varieties

To translate Hecke operators defined in a geometric manner into group-theoretic operators (in a more automorphic way), we introduce here the Hilbert modular Shimura varieties. We give an exposition of more general Shimura varieties later in Chapters 6 and 7. Let $h_{\mathbf{0}} : \mathbb{S} = \mathrm{Res}_{\mathbb{C}/\mathbb{R}} \mathbb{G}_m \to G_{/\mathbb{R}}$ be

the homomorphism of real algebraic groups sending $a + b\sqrt{-1} \in \mathbb{S}(\mathbb{R}) = \mathbb{C}^\times$ to the matrix $\left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right)$. We write X for the conjugacy class of $h_{\mathbf{0}}$ under $G(\mathbb{R})$ (with origin $\mathbf{0} = h_{\mathbf{0}}$). The group $G(\mathbb{R})$ acts on X from the left by conjugation. Since the centralizer of $h_{\mathbf{0}}$ is the product of the maximal compact subgroup of the identity-connected component $G(\mathbb{R})_+$ of the real Lie group $G(\mathbb{R})$ and its center $Z(\mathbb{R})$, the identity-connected component X^+ containing $\mathbf{0} = h_{\mathbf{0}}$ is isomorphic by $g(\mathbf{0}) \mapsto g(\mathbf{i})$ ($\mathbf{i} = (\sqrt{-1}, \ldots, \sqrt{-1})$) to the product $\mathfrak{Z} = \mathfrak{H}^I$. Here the action of $(g_\sigma)_{\sigma \in I} \in G(\mathbb{R})$ with $g_\sigma = \left(\begin{smallmatrix} a_\sigma & b_\sigma \\ c_\sigma & d_\sigma \end{smallmatrix}\right)$ on $\mathfrak{Z}$ is given by $z = (z_\sigma) \mapsto \left(\frac{a_\sigma z_\sigma + b_\sigma}{c_\sigma z_\sigma + d_\sigma}\right)$. Thus X is a finite disjoint union of the Hermitian symmetric domain isomorphic to $\mathfrak{Z}$, and for an arithmetic subgroup $\Gamma \subset G(\mathbb{Q})$, $\Gamma\backslash X$ is a finite disjoint union of connected Hilbert modular varieties.

The pair (G, X) satisfies Deligne's axioms for admitting Shimura varieties in [D2] 2.1.1 (see (D1–4), (SC), and (CT) in 7.2.1 in the text). The Shimura variety is a proalgebraic variety (i.e., a projective limit of algebraic varieties) defined over a canonical field of definition (called the *reflex field* of (G, X)) equipped with a scheme-theoretic action of $G(\mathbb{A}^{(\infty)})$ which characterizes the variety. The complex (pro-)analytic space of complex points of the Shimura variety is given by a projective limit under the inclusion relation of open compact subgroups of $G(\mathbb{A}^{(\infty)})$:

$$\begin{aligned} Sh(\mathbb{C}) = Sh(G, X)(\mathbb{C}) &= \varprojlim_K G(\mathbb{Q})\backslash \left(X \times G(\mathbb{A}^{(\infty)})\right)/K \\ &\overset{(*)}{=} G(\mathbb{Q})\backslash \left(X \times G(\mathbb{A}^{(\infty)})\right)/\overline{Z(\mathbb{Q})} \overset{(**)}{=} G(\mathbb{Q})_+\backslash \left(X^+ \times G(\mathbb{A}^{(\infty)})\right)/\overline{Z(\mathbb{Q})}, \end{aligned} \tag{4.21}$$

where $(\gamma, u) \in G(\mathbb{Q}) \times K$ acts on $(x, g) \in X \times G(\mathbb{A}^{(\infty)})$ by $\gamma(x, g)u = (\gamma_\infty(x), \gamma^{(\infty)} g u)$, and $\overline{Z(\mathbb{Q})}$ is the topological closure of the center $Z(\mathbb{Q})$ in $G(\mathbb{A}^{(\infty)})$ (see [EPE] Chapter II for a detailed discussion of the topology on adele groups). We leave for the reader the proof of the above identity $(*)$ in (4.21), which is not difficult (see [D2] 2.1.10 or [Mi] page 324 and Lemma 10.1). The identity $(**)$ follows from the fact that $G(\mathbb{Q})$ acts transitively on the set $\pi_0(X)$ of connected components of X (and $G(\mathbb{Q})_+$ is the stabilizer of X^+). On $Sh(G, X)(\mathbb{C})$, the group $G(\mathbb{A}^{(\infty)})$ acts by right multiplication $x \mapsto xg$. This proalgebraic variety has a unique canonical model $Sh(G, X)$ for $G = \mathrm{Res}_{F/\mathbb{Q}} GL(2)$ defined over the reflex field $\mathbb{Q}$ of (G, X), as we show later. In this section, assuming the representability of the classification functors introduced in Section 4.1.2, we review the construction of the model, emphasizing its automorphism group $\overline{\mathcal{E}}(G, X)$ (fixing a geometrically irreducible component). Strictly speaking, the group $\overline{\mathcal{E}}(G, X)$ we study is a subgroup of finite index in the full automorphism group, and the full automorphism group is a semi-direct product of $\overline{\mathcal{E}}(G, X)$ with the field automorphism group $\mathrm{Aut}(F)$. As is clear from Shimura's original construction of canonical models [Sh5] (see also [AAF] Chapter II) and the later treatment of Deligne

[D2], the full knowledge of $\overline{\mathcal{E}}(G,X)$ is almost equivalent to the existence of the canonical model itself.

4.2.1 Abelian Varieties up to Isogenies

Let $V = F^2$ be a column vector space, and put $V(\mathbb{A}^{(\infty)}) = V_{\mathbb{A}^{(\infty)}} := V \otimes_{\mathbb{Q}} \mathbb{A}^{(\infty)}$. We often write $F_{\mathbb{A}^{(\infty)}}$ for $F \otimes_{\mathbb{Q}} \mathbb{A}^{(\infty)}$, which is the finite part of the adele ring $F_{\mathbb{A}} = F \otimes_{\mathbb{Q}} \mathbb{A}$. Then $V(\mathbb{A}^{(\infty)})$ is an $F_{\mathbb{A}^{(\infty)}}$-free module of rank 2. We consider the fibered category $\mathcal{A}_F^{\mathbb{Q}}$ over $\mathbb{Q}$-SCH defined by the following data:

(Object) abelian schemes with real multiplication by O;
(Morphism) $\mathrm{Hom}_F^{\mathbb{Q}}(A,A') = \mathrm{Hom}_O(A,A') \otimes_{\mathbb{Z}} \mathbb{Q}$.

For an object $A_{/S}$, we take a geometric point $s \in S$, consider the Tate module $\mathcal{T}(A) = \mathcal{T}_s(A) = \varprojlim_N A[N](k(s))$, and define $V(A) = V_s(A) = \mathcal{T}(A) \otimes_{\mathbb{Z}} \mathbb{A}^{(\infty)}$. The module $V(A)$ is an $F_{\mathbb{A}^{(\infty)}}$-free module of rank 2 and has an $\widehat{O}$-stable lattice $\mathcal{T}(A)$, where $\widehat{O} = O \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}} = \prod_{\ell:\mathrm{prime}} O_\ell$.

A full level structure on A is an isomorphism $\eta : V(\mathbb{A}^{(\infty)}) \cong V_s(A)$ of $F_{\mathbb{A}^{(\infty)}}$-modules, picking a geometric point s in each connected component of S. For a closed subgroup $K \subset G(\mathbb{A}^{(\infty)})$, a level K-structure is the K-orbit $\overline{\eta} = \eta K$ of η for the right action $\eta \mapsto \eta \circ u$ $(u \in K)$. For many instances, we assume K to be open compact. Since $A[N]_{/S}$ is an étale finite group scheme, the algebrogeometric fundamental group $\pi_1(S,s)$ with base point s (see Section 4.4) acts on $A[N](k(s))$ and hence on $V_s(A)$. The level K-structure is defined over S if $\sigma \circ \overline{\eta} = \overline{\eta}$ for each $\sigma \in \pi_1(S,s)$. If the compatibility $\sigma \circ \overline{\eta} = \overline{\eta}$ is valid at one geometric point s for each connected component of S, it is valid for all $s \in S$ (see Section 6.4.1 for the proof of this fact given for Siegel modular varieties).

Two polarizations $\lambda, \lambda' : A \to {}^tA$ are said to be equivalent (written as $\lambda \sim \lambda'$) if $\lambda = a\lambda' = \lambda' \circ a$ for a totally positive $a \in F$. Here a is any fraction in $F_+^\times$ (not just an integer in O). Without introducing the category $\mathcal{A}_F^{\mathbb{Q}}$ up to isogeny, our notion of polarization classes does not make sense. The equivalence class of a polarization λ defined over S is written as $\overline{\lambda}$. As we have seen, if the class $\overline{\lambda}$ is defined over S, we can find a polarization $\lambda \in \overline{\lambda}$ really defined over S. We recall the following condition in Section 4.1.1 imposed on all triples $(A, \overline{\lambda}, \overline{\eta})_{/S}$ which is equivalent to (rm4) there.

(det) The characteristic polynomial of each $\alpha \in O$ on $Lie(A)$ over $\mathcal{O}_S$ is given by the image of $\prod_\sigma (t - \sigma(\alpha)) \in \mathbb{Z}[t]$ in $\mathcal{O}_S[t]$, where σ runs over all embeddings of F into $\overline{\mathbb{Q}}$.

We consider the following functor from $SCH_{/\mathbb{Q}}$ into $SETS$,

$$\mathcal{E}_K^{(\emptyset)}(S) = \left[(A, \overline{\lambda}, \overline{\eta})_{/S} \text{ with (det)} \;\middle|\; \overline{\eta} \text{ is a level } K\text{-structure}\right].$$

Here $[\ ] = \{\ \}/\cong$ indicates the set of isomorphism classes of the objects defined over S in the brackets. An F-linear morphism $\phi \in \mathrm{Hom}_F^{\mathbb{Q}}(A,A')$ is an

isomorphism between triples $(A, \overline{\lambda}, \overline{\eta})_{/S}$ and $(A', \overline{\lambda}', \overline{\eta}')_{/S}$ if it is compatible with all data; that is,

$$\phi \circ \overline{\eta} = \overline{\eta}' \text{ and } {}^t\phi \circ \overline{\lambda} = \overline{\lambda}' \circ \phi.$$

Equip $V = F^2$ with an alternating form $\Lambda : V \wedge_F V \cong F$ given by $(x, y) = {}^t x J_1 y$ for $J_1 = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$. We define a $\mathbb{Q}$-alternating pairing $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{Q}$ by $\mathrm{Tr}_{F/\mathbb{Q}} \circ \Lambda$. Suppose the point $s \in S$ is a complex point $s \in S(\mathbb{C})$; so, we have the Betti homology group $H_1(A, \mathbb{Q}) := H_1(A(k(s)), \mathbb{Q})$. Then the polarization $\lambda : A \to {}^t A$ induces a nondegenerate F-Hermitian alternating pairing $E_\lambda : \bigwedge^2 H_1(A, \mathbb{Q}) \to \mathbb{Q}$ (the Riemann form; see [ABV] Sections 1 and 20 and also the proof of Theorem 8.9 in the text). Here the word: "F-Hermitian" means $E_\lambda(\alpha x, y) = E_\lambda(x, \alpha y)$ for all $\alpha \in F$. We write $e_\lambda : V \wedge_F V \cong F$ for a unique alternating form satisfying $\mathrm{Tr}_{F/\mathbb{Q}} \circ e_\lambda = E_\lambda$. The Hodge decomposition: $H^1(A, \mathbb{C}) = H^0(A(k(s)), \Omega^{an}_{A/\mathbb{C}}) \oplus H^0(A(k(s)), \overline{\Omega}^{an}_{A/\mathbb{C}})$ induces, by Poincaré duality, an embedding $h = h_A : \mathbb{C}^\times = \mathbb{S}(\mathbb{R}) \to \mathrm{Aut}_F(H_1(A, \mathbb{R}))$ such that

1. $h(z)\omega = z\omega$ for all $\omega \in \mathrm{Hom}_{\mathbb{C}}(H^0(A(k(s)), \Omega_{A/\mathbb{C}}), \mathbb{C})$ (and $h(z)\overline{\omega} = \overline{z\omega}$);
2. $E_\lambda(x, h(\sqrt{-1})y)$ is a positive definite Hermitian form on $V_{\mathbb{R}} = V \otimes_{\mathbb{Q}} \mathbb{R}$ under the complex structure given by h.

In the above definition of $\mathcal{E}^{(\emptyset)}_K$, a condition usually required is missing:

(pol) There exists an F-linear isomorphism $f : V \cong H_1(A, \mathbb{Q})$ such that $f^{-1} \circ h_A \circ f$ is a conjugate of $h_{\mathbf{0}}$ under $G(\mathbb{R})$, $f \equiv \eta \mod K$ under the canonical isomorphism $V_s(A) \cong H_1(A, \mathbb{A}^{(\infty)}) = H_1(A, \mathbb{Q}) \otimes_{\mathbb{Q}} \mathbb{A}^{(\infty)}$ and $e_\lambda(f(x), f(y)) = \alpha \cdot \Lambda(x \wedge y)$ for some $\alpha \in F^\times$.

For every nondegenerate F-Hermitian alternating form E on V and an open compact subgroup $K \subset G(\mathbb{A}^{(\infty)})$, we can find $\alpha \in G(\mathbb{Q})$ such that $E(x, y) \equiv \langle \alpha x, \alpha y \rangle \mod K$ (the strong approximation theorem). Since $G(\mathbb{R})$ is the full group of $F_{\mathbb{R}}$-linear automorphisms of $V_{\mathbb{R}}$, $f^{-1} \circ h_A \circ f$ is always conjugate to $h_{\mathbf{0}}$. Thus this condition is redundant; so, we ignore it.

By [Sh3] and [D1] 4.16–21, the canonical model $Sh(G, X)_{/\mathbb{Q}}$ represents the functor $\mathcal{E}^{(\emptyset)}_1$ over $\mathbb{Q}$ for the trivial group 1 made of the identity element. Through the action of $G(\mathbb{A}^{(\infty)})$ on $F^2_{\mathbb{A}^{(\infty)}}$, $g \in G(\mathbb{A}^{(\infty)})$ acts on the level structure by $\eta \mapsto \eta \circ g$ and hence on the variety $Sh(G, X)$ from the right. If K is open and sufficiently small (so that $\mathrm{Aut}((A, \overline{\lambda}, \overline{\eta})_{/S}) = \{1\}$ for all test objects $(A, \overline{\lambda}, \overline{\eta})_{/S}$), $Sh_K(G, X) := (Sh(G, X)/K)_{/\mathbb{Q}}$ (whose complex points are given by $G(\mathbb{Q})\backslash(X \times G(\mathbb{A}^{(\infty)}))/K$) represents $\mathcal{E}^{(\emptyset)}_K$ over $\mathbb{Q}$.

We now give a brief outline of the proof of the representability, reducing it to the representability of the functors similar to $\mathcal{E}_{\Gamma(N)}$ in Section 4.1.2. By shrinking K, we may assume that $\det(K) \cap O^\times_+ \subset (K \cap Z(\mathbb{Z}))^2$. This is to guarantee that the images of $gKg^{-1} \cap G_1(\mathbb{Q})$ and $gKg^{-1} \cap G(\mathbb{Q})^+$ in $PG(\mathbb{Q})$ are equal; so, $Sh_K(\mathbb{C})$ can be embedded into $Sh_{K_1}(\mathbb{C})$ for $K_1 = G_1(\mathbb{A}^{(\infty)}) \cap K$, because the moduli problem with respect to K_1 is neat without having any

nontrivial automorphisms. Let $L \subset V$ be an O-lattice. We may assume that $L = \mathfrak{a}^* \oplus \mathfrak{b}$ for a pair $(\mathfrak{a}, \mathfrak{b})$ of two fractional ideals. We define the polarization ideal $\mathfrak{c}$ by $F \supset \mathfrak{c}^* = \Lambda(L \wedge L)$. For each point $h_z \in X$, we have a unique point $z \in (\mathbb{C} - \mathbb{R})^I$ fixed by $h_z(\mathbb{C}^\times)$ (in this way, we identify $\mathfrak{Z}$ with the connected component X^+ of X containing h_0). By changing the identification $V = F^2$, we may assume that $z \in X^+ = \mathfrak{Z}$. The action of $h_z(\mathbb{C}^\times)$ on $V_{\mathbb{R}} = V \otimes_{\mathbb{Q}} \mathbb{R}$ gives a structure of a complex vector space of dimension $g = [F : \mathbb{Q}]$ on $V_{\mathbb{R}}$; that is, $V_{\mathbb{R}} = \mathbb{C}^I$ via $(a, b) \mapsto -a + bz = (a, b)J_1 \cdot {}^t(z, 1)$ for $J_1 = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$. Then $L \subset V_{\mathbb{R}}$ gives rise to the lattice $\mathcal{L}_z$, and Λ induces the $\mathfrak{c}$-polarization λ_z, and we recover the data $(\mathcal{L}_z, \lambda_z)$ introduced in Section 4.1.3. Set $\widehat{L} = L \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}} \subset V_{\mathbb{A}}^{(\infty)} = V \otimes_{\mathbb{Q}} \mathbb{A}^{(\infty)}$, and define an abelian variety $A_{z/\mathbb{C}}$ by $A_z(\mathbb{C}) = \mathbb{C}^I/\mathcal{L}_z$. Then we have $\mathcal{T}(A_z) = \widehat{L}$, which induces $\eta_z : V_{\mathbb{A}^{(\infty)}} \cong V(A_z)$ and gives rise to a level N-structure $\phi_N : N^{-1}L/L \cong A_z[N]$ for any $N > 0$.

Let $Cl(K) = F_{\mathbb{A}^{(\infty)}}^\times / \det(K) F_+^\times$, which is a finite ideal class group. We fix a complete representative set for $Cl(K)$ in fractional ideals $\mathfrak{c}$ of F. We choose an O-lattice $L_{\mathfrak{c}} = \mathfrak{a}^* \oplus \mathfrak{b} \subset V$ as above with $\Lambda(L_{\mathfrak{c}} \wedge L_{\mathfrak{c}}) = \mathfrak{c}^*$ ($\Leftrightarrow \mathfrak{a}\mathfrak{b}^{-1} = \mathfrak{c}$).

We pick an element $(A, \overline{\lambda}, \overline{\eta})_{/S} \in \mathcal{E}_K^{(\emptyset)}(S)$. Let $\widehat{\mathbb{Z}}(1) = \varprojlim_N \mu_N$ and identify it with $\widehat{\mathbb{Z}}$ by choosing a primitive Nth root $\zeta_N = \mathbf{e}(\frac{1}{N})$ in $\overline{\mathbb{Q}}$. The duality pairing $e : \mathcal{T}(A) \wedge_{\widehat{\mathbb{Z}}} \mathcal{T}({}^tA) \twoheadrightarrow \widehat{\mathbb{Z}}(1) = \widehat{\mathbb{Z}}$ (as in (P1–3) of Section 8.2.3) is O-Hermitian (because $O \subset \mathrm{End}(A_{/S})$ is invariant under the Rosati involution). Thus we have a unique O-linear alternating form $e_O(A) : \mathcal{T}(A) \wedge \mathcal{T}({}^tA) \cong \widehat{O}^*$ such that $e = \mathrm{Tr}_{F/\mathbb{Q}} \circ e_O(A)$. If one makes a choice of a polarization $\lambda \in \overline{\lambda}$, the duality pairing e_O composed with $\lambda : \mathcal{T}(A) \to \mathcal{T}({}^tA)$ (and pulled back by η) induces an alternating pairing $\Lambda_A : \eta^{-1}(\mathcal{T}(A)) \wedge \eta^{-1}(\mathcal{T}(A)) \cong \widehat{\mathfrak{c}}^* \subset F_{\mathbb{A}}^{(\infty)}$ with $\Lambda_A \equiv \Lambda \mod \det(K)$. This means that $\Lambda = \Lambda_A \circ g$ for $g \in G(\mathbb{A}^{(\infty)})$ with $\det(g) \in \det(K)$. Under this condition, the fractional ideal $\mathfrak{c}$ is uniquely determined in the complete representative set for $Cl(K)$ we have chosen, because $\overline{\lambda}$ is a $F_+^\times$-orbit of λ and $\overline{\eta}$ is the K-orbit of η.

Case 1: We first assume that $\eta(\widehat{L}_{\mathfrak{c}})$ contains $\mathcal{T}(A)$. We can naturally identify $C = \eta(\widehat{L}_{\mathfrak{c}})/\mathcal{T}(A)$ with a subgroup of A. For each geometric point $s \in S$ and $\sigma \in \pi_1(S, s)$, we have $\sigma \circ \eta = \eta \circ \rho(\sigma)$ for $\rho(\sigma) \in K$, and for $x = \eta(v)$ mod $\mathcal{T}(A)$ with $v \in \widehat{L}_{\mathfrak{c}}$, we have $\sigma(x) = \sigma(\eta(v)) = \eta(\rho(\sigma)v)$ and $\rho(\sigma)v \in \widehat{L}_{\mathfrak{c}}$, because $\widehat{L}_{\mathfrak{c}}$ is stable under K. This implies that the subgroup C is stable under $\pi_1(S, s)$ for all geometric points $s \in S$, and hence C is a subgroup scheme of A defined over S. We make the quotient $A' = A/C$ which is an abelian scheme over S (cf. [ABV] Section 12). For the quotient map $\pi : A \to A'$, we define $\lambda_{A'} = {}^t\pi^{-1} \circ \lambda \circ \pi^{-1}$. By definition, $A'_{/S}$ has $\pi_1(S, s)$-equivariant exact sequence $0 \to \mathcal{T}(A) \to \mathcal{T}(A') \to C \to 0$ and hence $V(A') = V(A)$, which induces $\eta : V_{\mathbb{A}}^{(\infty)} \cong V(A) = V(A')$ with the exact identity $\eta(\widehat{L}_{\mathfrak{c}}) = \mathcal{T}(A')$. The polarization of A induces an equivalent polarization $\lambda_{A'}$ on A', and we have $i : F \hookrightarrow \mathrm{End}_F^{\mathbb{Q}}(A) = \mathrm{End}_F^{\mathbb{Q}}(A')$. Thus as elements in $\mathcal{E}_K^{(\emptyset)}(S)$, $(A, \overline{\lambda}, \overline{\eta})_{/S}$ and $(A', \overline{\lambda}_{A'}, \overline{\eta})_{/S}$ are the same.

Case 2: When $\eta(\widehat{L}_{\mathfrak{c}}) \subset \mathcal{T}(A)$, by the canonical identification of the algebraic fundamental group $\pi_1(A, \mathbf{0}) = \mathcal{T}(A)$ (see [ABV] Section 18), we have a unique abelian variety $A'_{/S}$ with an F-linear étale isogeny $\pi : A' \twoheadrightarrow A$ with $\mathcal{T}(A') = \eta(\widehat{L}_{\mathfrak{c}})$. Then in the same manner as above, we create $(A', \overline{\lambda}_{A'}, \overline{\eta})_{/S}$ giving rise to the same class of $(A, \overline{\lambda}, \overline{\eta})$. In particular, $\lambda_{A'} = {}^t\pi \circ \lambda \circ \pi$.

Case 3: If $\eta(\widehat{L}_{\mathfrak{c}})$ does not have either inclusion relation with $\mathcal{T}(A)$, we take O-lattices L' with $\widehat{L}' = \eta^{-1}(\mathcal{T}(A)) + \widehat{L}_{\mathfrak{c}}$ and L'' with $\widehat{L}'' = \eta^{-1}(\mathcal{T}(A)) \cap \widehat{L}_{\mathfrak{c}}$. Starting from A, first create $(A'', \overline{\lambda}'', \overline{\eta}'')_{/S}$ in the class of $(A, \overline{\lambda}, \overline{\eta})$ with $\mathcal{T}(A'') = \eta(\widehat{L}')$. Then from $(A'', \overline{\lambda}'', \overline{\eta}'')_{/S}$, we create $(A', \overline{\lambda}_{A'}, \overline{\eta})_{/S}$ with $\mathcal{T}(A') = \eta(\widehat{L}_{\mathfrak{c}})$.

By the above argument, in each class of $(A, \overline{\lambda}, \overline{\eta})_{/S}$, we can find a unique $(A', \overline{\lambda}_{A'}, \overline{\eta})_{/S}$ with $\eta(\widehat{L}_{\mathfrak{c}}) = \mathcal{T}(A')$. By our choice of $L_{\mathfrak{c}}$ and $\mathfrak{c}$: $\Lambda(L_{\mathfrak{c}} \wedge L_{\mathfrak{c}}) = \mathfrak{c}^*$, actually only Case 3 occurs, because Λ and $\Lambda_{A'}$ are proportional up to units in $\det(K)$. So we may assume that $\Lambda = \Lambda_{A'}$, because the level structure is defined modulo K.

In any case, $\overline{\lambda}_{A'}$ still contains a unique $\mathfrak{c}$-polarization $\lambda_{A'}$ made from λ. Thus once we have chosen a $\mathfrak{c}$-polarization in $\overline{\lambda}$ for each member $(A, \overline{\lambda}, \overline{\eta})$, we have a unique triple $(A', \lambda_{A'}, \overline{\eta})_{/S}$ with $\mathfrak{c}$-polarization $\lambda_{A'}$. If two such choices are isogenous, the isogeny between them has to be an isomorphism keeping the polarization. We have proven $\mathcal{E}_K^{(\emptyset)}(S) \cong \mathcal{E}'_K(S) := \bigsqcup_{\mathfrak{c} \in Cl(K)} \mathcal{E}'_{K,\mathfrak{c}}(S)$, where $\mathfrak{c}$ runs over the ideal classes in $Cl(K) = {F_{\mathbb{A}}^{(\infty)}}^{\times}/F_+^{\times} \det(K)$, and

$$\mathcal{E}'_{K,\mathfrak{c}}(S) = \left\{ (A, \lambda, \overline{\eta})_{/S} \text{ with (det)} \;\middle|\; \eta(\widehat{L}_{\mathfrak{c}}) = \mathcal{T}(A) \;\text{ and }\; \mathfrak{c}(\lambda) = \mathfrak{c} \right\} / \cong .$$

Here $\cong$ is the isomorphism (not an isogeny) for a chosen polarization integral over the fixed lattice $L_{\mathfrak{c}}$ in the class of $\overline{\lambda}$ (in other words, λ induces a fixed alternating form on the V integral over $L_{\mathfrak{c}}$ up to units in $\det(K)$).

We still need to worry that one might not be able to choose λ in $\overline{\lambda}$ functorially (depending on S). In other words, the above bijection $\mathcal{E}_K^{(\emptyset)}(S) \cong \mathcal{E}'_K(S)$ may not be induced from an isomorphism of functors. Under the condition $\det(K) \cap O_+^{\times} \subset (K \cap Z(\mathbb{Z}))^2$, this ambiguity of the choice of $\lambda_{A'}$ does not matter. Indeed, since $\Lambda_{A'} \equiv \Lambda \mod \det(K)$, any $\mathfrak{c}$-polarization λ' in $\overline{\lambda}_{A'}$ inducing $\Lambda' \equiv \Lambda \mod \det(K)$ is given by $\epsilon\lambda_{A'}$ for $\epsilon \in \det(K) \cap O_+^{\times}$. By $\det(K) \cap O_+^{\times} \subset (K \cap Z(\mathbb{Z}))^2$, take $\varepsilon \in K \cap Z(\mathbb{Z})$ with $\varepsilon^2 = \epsilon$. Writing $i : O \hookrightarrow \mathrm{End}(A'_{/S})$ for the O-multiplication, we have

$$(A', \epsilon\lambda_{A'}, \overline{\eta}) = (A', {}^t i(\varepsilon) \circ \lambda_{A'} \circ i(\varepsilon) = \lambda', \overline{\eta}) = (A', \lambda_{A'}, \varepsilon\overline{\eta}) = (A', \lambda_{A'}, \overline{\eta})$$

in $\mathcal{E}'_K(S)$, since we have $\varepsilon\overline{\eta} = \overline{\eta}$ because $\varepsilon \in K$. Thus the choice of λ also does not matter in the target $\mathcal{E}'_K$, and the above bijection is indeed induced by an isomorphism of functors if $\det(K) \cap O_+^{\times} \subset (K \cap Z(\mathbb{Z}))^2$.

If we take $L = \mathfrak{d}^{-1} \oplus O \subset V$, then $L \wedge L \cong O^*$ by $(a, b) \wedge (a', b') \mapsto a'b - ab'$. For this choice, the polarization ideal is trivial; that is, $\mathfrak{c} = O$. If further, we

identify μ_N with $\mathbb{Z}/N\mathbb{Z}$ by choosing a primitive Nth root $\zeta = \zeta_N$ of unity in $\mathbb{Q}[\mu_N]$, having $\overline{\eta}$ is equivalent to having

$$\phi_N : (\mathfrak{d}^{-1} \otimes \mu_N) \times (O \otimes \mathbb{Z}/N\mathbb{Z}) \cong N^{-1}L/L \overset{\overline{\eta}}{\cong} A[N].$$

Take the open compact subgroup $S(N)$ in (4.6), a conjugate of $\widehat{\Gamma}(N)$, stabilizing the lattice $L = \mathfrak{d}^{-1} \oplus O$. Over $\mathbb{Q}[\mu_N]$, the component of $Sh_{S(N)}(G, X)$ corresponding to $\mathcal{E}'_{S(N),O}$ (which corresponds in turn to $G(\mathbb{Q})S(N)G(\mathbb{R})^+$ in (4.7)) represents the now familiar functor:

$$\mathcal{E}_{\Gamma(N)} \text{ over } \mathbb{Q}[\mu_N]\text{-}SCH \text{ if } L = \mathfrak{d}^{-1} \oplus O \subset V \text{ with } \mathfrak{c} = O. \tag{4.22}$$

The choice ζ gives rise to the identification $\mathbb{Q}[\mu_N] = \mathbb{Q}[X]/(\Phi_N(X))$ with $\mathbb{Q}[\zeta]$ for the cyclotomic polynomial $\Phi_N(X)$ in 2.2.1, and an automorphism $\sigma \in \mathrm{Gal}(\mathbb{Q}[\zeta]/\mathbb{Q})$ changes the identification by $\zeta \mapsto \zeta^\sigma$, whose action is induced by $\phi_N \mapsto \phi_N \circ \left(\begin{smallmatrix} c & 0 \\ 0 & 1 \end{smallmatrix}\right)$ for a unit $c \in \widehat{\mathbb{Z}}^\times$ such that $\zeta^\sigma = \zeta^c$. In other words, the action of $\left(\begin{smallmatrix} c & 0 \\ 0 & 1 \end{smallmatrix}\right) \in G(\mathbb{A}^{(\infty)})$ on $Sh(G, X)_{/\mathbb{Q}}$ brings $\mathfrak{M}(O, \Gamma(N))_{/\mathbb{Q}[\zeta]} \subset Sh_{S(N)}(G, X)$ to its σ-conjugate $\mathfrak{M}(O, \Gamma(N))^\sigma_{/\mathbb{Q}[\zeta]}$ in $Sh_{S(N)}(G, X)_{/\mathbb{Q}}$.

To accommodate all other components $G(\mathbb{Q})\left(\begin{smallmatrix} c & 0 \\ 0 & 1 \end{smallmatrix}\right)S(N) \subset G(\mathbb{A}^{(\infty)})$ is to change the lattice L to $L_c = \left(\widehat{L} \cdot \left(\begin{smallmatrix} c & 0 \\ 0 & 1 \end{smallmatrix}\right)^{-1}\right) \cap V$. Thus we have

$$\mathcal{E}'_{S(N)} = \bigsqcup_{\mathfrak{c} \in Cl^+_F(N)} \mathcal{E}_{\mathfrak{c},\Gamma(N)} \text{ over } \mathbb{Q}[\mu_N]\text{-}SCH, \tag{4.23}$$

which implies

$$Sh_{S(N)}(G, X)_{/\mathbb{Q}[\mu_N]} = \bigsqcup_{\mathfrak{c} \in Cl^+_F(N)} \mathfrak{M}(\mathfrak{c}, \Gamma(N)) \text{ over } \mathbb{Q}[\mu_N]. \tag{4.24}$$

Note that the identity over $\mathbb{Q}$ is wrong, because, over $\mathbb{Q}$, the group scheme $(\mu_N \otimes \mathfrak{d}^{-1}) \times (O/NO)$ is not isomorphic to the constant group $(\mathfrak{d}^{-1}/N\mathfrak{d}^{-1}) \times (O/NO)$. In any case, by the Galois action on $\mathfrak{M}(\mathfrak{c}, \Gamma(N))_{/\mathbb{Q}[\mu_N]}$ described just below (4.22), we can descend the right-hand side to the base field $\mathbb{Q}$ to obtain the model $Sh_{S(N)}(G, X)_{/\mathbb{Q}}$, because $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ is quasi-projective as we already observed (see [GME] Example 1.11.1 for the descent). Thus if $F = \mathbb{Q}$, $Sh_{S(N)}(GL(2), X)$ is irreducible and is isomorphic to the scheme $\mathfrak{M}(1, \Gamma(N)) \otimes_{\mathbb{Q}} \mathbb{Q}[\mu_N]$ regarded as defined over $\mathbb{Q}$, but not geometrically irreducible (whose geometrically irreducible components are isomorphic to $Y(N) = \mathfrak{M}(1, \Gamma(N))$ over $\overline{\mathbb{Q}}$). For $F \neq \mathbb{Q}$, $Sh_{S(N)}(G, X)$ may not even be irreducible over $\mathbb{Q}$, and the picture is not very straightforward (since it involves an explicit description of the action of $G(\mathbb{A}^{(\infty)})$ on the set $\pi_0(Sh(G, X)_{/\overline{\mathbb{Q}}})$ of connected components).

Take a sufficiently large integer N so that $S(N) \subset K$. Then the coarse moduli scheme representing $\mathcal{E}'_K$ is given by the quotient $Sh_{S(N)}(G, X)/K$.

Here $k \in K$ acts through its finite quotient $K/S(N)$ by $\overline{\eta} \mapsto \overline{\eta} \circ k$. As before, if K is sufficiently small so that $(A, \lambda, \overline{\eta})$ is rigid for all test objects, the scheme represents the functor. Although $\mathcal{E}'_K$ can be representable over the integral base $\mathbb{Z}\left[\frac{1}{N}\right]$, the process bringing $\mathcal{E}_K^{(\emptyset)}$ isomorphically to $\mathcal{E}'_K$ fully uses the fact that the group scheme C has order invertible in $\mathbb{Q}$; so, the functor $\mathcal{E}_K^{(\emptyset)}$ is representable over $\mathbb{Q}$ but not on the integral base. We write the scheme representing $\mathcal{E}_K^{(\emptyset)}$ as $Sh_K(G,X)$ and $Sh(G,X) = \varprojlim_K Sh_K(G,X)$, which is defined over $\mathbb{Q}$. Then we recover (4.21) from (4.7).

To construct integral models of Shimura varieties, we use the following variant (due to Kottwitz [Ko]) of the functor $\mathcal{E}_K^{(\emptyset)}$. We fix a proper subset Σ of rational primes. This concerns an open-compact subgroup K *maximal at* Σ (i.e., $K = G(\mathbb{Z}_\Sigma) \times K^{(\Sigma)}$), where $\mathbb{Z}_\Sigma = \prod_{p\in\Sigma} \mathbb{Z}_p$ and $O_\Sigma = O \otimes_{\mathbb{Z}} \mathbb{Z}_\Sigma$. As usual, we have written $K^{(\Sigma)} = \{x \in K | x_p = 1 \text{ for all } p \in \Sigma\}$. Recall $\mathbb{A}^{(\Sigma\infty)} = \{x \in \mathbb{A} | x_v = 0 \text{ for all } v \in \Sigma \cup \{\infty\}\}$. We identify the multiplicative group $\mathbb{A}^{(\Sigma\infty)^\times}$ with $\{x \in \mathbb{A}^\times | x_v = 1 \text{ for all } v \in \Sigma \cup \{\infty\}\}$.

We consider the following fibered category $\mathcal{A}_F^{(\Sigma)}$ over $\mathbb{Z}_{(\Sigma)}$-schemes:

(Object) abelian schemes with real multiplication by O;
(Morphism) We define $\mathrm{Hom}_{\mathcal{A}_F^{(\Sigma)}}(A, A') = \mathrm{Hom}_{\mathcal{A}_F}(A, A') \otimes_{\mathbb{Z}} \mathbb{Z}_{(\Sigma)}$, where

$$\mathbb{Z}_{(\Sigma)} = \left\{ \frac{a}{b} \middle| b\mathbb{Z} + p\mathbb{Z} = \mathbb{Z} \text{ for all } p \in \Sigma \right\}.$$

This means that to classify test objects, we now allow only isogenies with degree prime to all primes in Σ (called "prime-to-Σ isogenies"), and the degree of the polarization λ is supposed to be also prime to Σ. Two polarizations are equivalent if $\lambda = a\lambda' = \lambda' \circ a$ for a totally positive a prime to Σ.

Fix a prime $p \in \Sigma$ and an O-lattice $L \subset V = F^2$ with $\Lambda(L \wedge L) = \mathfrak{c}^*$, and assume self O_p-duality of $L_p = L \otimes_{\mathbb{Z}} \mathbb{Z}_p$ under the O-Hermitian alternating pairing $\Lambda : V \wedge V \cong F$. Consider test objects $(A, \overline{\lambda}, \overline{\eta}^{(\Sigma)})_{/S}$. Here $\eta^{(\Sigma)} : V(\mathbb{A}^{(\Sigma\infty)}) = V \otimes_{\mathbb{Q}} \mathbb{A}^{(\Sigma\infty)} \cong V^{(\Sigma)}(A) = \mathcal{T}(A) \otimes_{\mathbb{Z}} \mathbb{A}^{(\Sigma\infty)}$ and $\lambda \in \overline{\lambda}$ are supposed to satisfy the following requirement, $\mathcal{T}^{(\Sigma)}(A) \wedge \mathcal{T}^{(\Sigma)}(A) \xrightarrow{e_\lambda} F_{\mathbb{A}}^{(\Sigma\infty)}$ is proportional to $\Lambda : V \wedge V \cong F$ up to scalars in $(F \otimes \mathbb{A}^{(\Sigma\infty)})^\times$. We write the $K^{(\Sigma)}$-orbit of $\eta^{(\Sigma)}$ as $\overline{\eta}^{(\Sigma)}$. Then we consider the following functor from $\mathbb{Z}_{(\Sigma)}$-schemes into $SETS$.

$$\mathcal{E}_K^{(\Sigma)}(S) = \left[(A, \overline{\lambda}, \overline{\eta}^{(\Sigma)})_{/S} \text{ with (det)} \right]. \tag{4.25}$$

Since all nondegenerate F_Σ-linear alternating forms on $V_\Sigma = V \otimes_{\mathbb{Q}} \mathbb{Q}_\Sigma$ are isomorphic (see Section 7.1.5) and there is only one isomorphism class of O_Σ-lattices in V_Σ, if S is a $\mathbb{Z}_{(p)}$-scheme, we can extend $\eta^{(\Sigma)} : V(\mathbb{A}^{(\Sigma\infty)}) \to V^{(\Sigma)}(A)$ to $\eta^{(p)} : V(\mathbb{A}^{(p\infty)}) \to V^{(p)}(A)$ (keeping rationality over S) so that $\eta^{(p)}$ sends the alternation form $\Lambda : V \wedge V \cong F$ to e_λ up to scalars in $(F_{\mathbb{A}}^{(p\infty)})^\times$ and $\eta_{\Sigma-\{p\}}(L_{\Sigma-\{p\}}) = \mathcal{T}_{\Sigma-\{p\}}(A)$. Indeed, defining $\rho : \pi_1(S,s) \to$

$G(\mathbb{Q}_{\Sigma-\{p\}})$ by $\sigma \circ \eta'_{\Sigma-\{p\}} = \eta'_{\Sigma-\{p\}} \circ \rho(\sigma)$ for a chosen identification $\eta'_{\Sigma-\{p\}} : V \otimes_{\Sigma-\{p\}} \mathbb{Q}_\Sigma \cong \mathcal{T}_{\Sigma-\{p\}}(A) \otimes_{\mathbb{Z}} \mathbb{Q}_{\Sigma-\{p\}}$, the image of ρ is in a maximal compact subgroup of $G(\mathbb{Q}_{\Sigma-\{p\}})$, which is a conjugate of $G(\mathbb{Z}_{\Sigma-\{p\}})$. By conjugating back, we may assume that $\eta_{\Sigma-\{p\}}$ is defined over S. By the argument similar to Steps 1-3, we may further assume $\eta_{\Sigma-\{p\}}(L_{\Sigma-\{p\}}) = \mathcal{T}_{\Sigma-\{p\}}(A)$. This extension $\eta^{(p)}$ is unique modulo $G(\mathbb{Z}_{\Sigma-\{p\}})$. In other words, as long as K is maximal at Σ, for each test object $(A, \overline{\lambda}, \overline{\eta}^{(\Sigma)})_{/S}$ for a $\mathbb{Z}_{(p)}$-scheme S, we have a unique test object $(A, \overline{\lambda}, \overline{\eta}^{(p)})_{/S}$, where $\overline{\eta}^{(p)}$ is the $K^{(p)}$-orbit of $\eta^{(p)}$, and we have the following expression on $\mathbb{Z}_{(p)}$-SCH.

$$\mathcal{E}_K^{(\Sigma)}(S) = \left[(A, \overline{\lambda}, \overline{\eta}^{(p)})_{/S} \text{ with (det)}\right]. \tag{4.26}$$

Let $O^\times_{(\Sigma)+} = O^\times_\Sigma \cap F^\times_+$. As long as Σ is finite and K is maximal at Σ, we can identify $Cl(K) = {F_{\mathbb{A}}^{(\infty)}}^\times / F^\times_+ \det(K)$ with ${F_{\mathbb{A}}^{(\Sigma\infty)}}^\times / O^\times_{(\Sigma)+} \det(K^{(\Sigma)})$. Thus we may choose the representative set $\{\mathfrak{c}\}$ prime to p (and hence we may assume the self-duality of L at p). Assuming that Σ is finite, by the same process as bringing $\mathcal{E}_K^{(\emptyset)}$ isomorphically to $\mathcal{E}'_{K/\mathbb{Q}}$, the functor is equivalent to $\mathcal{E}'_{K/\mathbb{Z}_{(p)}}$ defined over $\mathbb{Z}_{(p)}$-SCH; so, it is representable over $\mathbb{Z}_{(p)}$, giving a canonical model $Sh_K^{(\Sigma)}(G, X)_{/\mathbb{Z}_{(p)}}$ of $Sh_K(G, X)_{/\mathbb{Q}}$ over $\mathbb{Z}_{(p)}$. The functor $\mathcal{E}'_{K/\mathbb{Z}_{(p)}}$ is a disjoint union of functors $\mathcal{E}'_{K,\mathfrak{c}}$ indexed by $\mathfrak{c} \in Cl(K)$, where

$$\mathcal{E}'_{K,\mathfrak{c}}(S) = \left[(A, \lambda, \overline{\eta}^{(p)})_{/S} \text{ with (det) } \middle| \eta^{(p)}(\widehat{L}) = \mathcal{T}^{(p)}(A),\ \mathfrak{c}(\lambda) = \mathfrak{c}\right].$$

By $\mathfrak{c}(\lambda) = \mathfrak{c}$, the proportionality of Λ and e_λ and $\Lambda(L_p \wedge L_p) = O_p$, we must have $\Lambda(L \wedge L) = \mathfrak{c}^*$.

We now move around p inside Σ. We have reached the functor $\mathcal{E}'_{K/\mathbb{Z}_{(p)}}$ defined over $\mathbb{Z}_{(p)}$-SCH. Indeed, the subgroup C in the argument has order prime to p for any given $p \in \Sigma$ because we removed the Σ-components from $V(A)$; so, the process making the quotient can be performed even over $\mathbb{Z}_{(\Sigma)}$ (as long as $|\Sigma| < \infty$). Since $\mathcal{E}'_K$ is well-defined over $\mathbb{Z}_{(\Sigma)}$, moving around p, we find $\mathcal{E}_K^{(\Sigma)}$ is isomorphic to $\mathcal{E}'_K$ over $\mathbb{Z}_{(\Sigma)}$, and $\mathcal{E}_K^{(\Sigma)}$ is represented by a scheme $Sh_K^{(\Sigma)}$ over $\mathbb{Z}_{(\Sigma)}$. If Σ is not finite, we take an increasing sequence of finite subsets Σ_i with $\Sigma = \bigcup_i \Sigma_i$, and we have $Sh^{(\Sigma)}_{K/\mathbb{Z}_{(\Sigma)}} = \varinjlim_i Sh^{(\Sigma_i)}_{K/\mathbb{Z}_{(\Sigma_i)}}$; so, finiteness of Σ is not necessary to have an integral model over $\mathbb{Z}_{(\Sigma)}$ as long as K is maximal at Σ.

A subtle point is to relate $Sh^{(\Sigma)}_{/\mathbb{Z}_{(\Sigma)}}$ to $Sh_{/\mathbb{Q}}$. Since the above process of identifying functors $\mathcal{E}^{(\emptyset)}_{S(N)}$ and $\mathcal{E}^{(\Sigma)}_{S(N)}$ ends with the functor $\mathcal{E}'_{S(N)}$ independent of Σ (though their domain $\mathbb{Z}_{(\Sigma)}$-SCH depends on Σ) as long as N is prime to Σ, we conclude that $Sh^{(\Sigma)} \otimes_{\mathbb{Z}_{(\Sigma)}} \mathbb{Q} \cong (Sh/G(\mathbb{Z}_\Sigma))$. Since $\widehat{L}' \cong \widehat{L}$ for any O-lattice $L' \subset V$, it is essential to allow all O-isomorphism classes of O-lattices L' to define $\mathcal{E}'_{S(N)}$, because in the definition of $\mathcal{E}^{(\Sigma)}$, only $\widehat{L}^{(\Sigma)}$ is

specified (which does not determine the isomorphism class of L if the class group of F is nontrivial). This problem is more acute at p because over $\mathbb{Z}_{(p)}$, $\mathcal{T}_p(A)$ does not determine L_p. Indeed the p-adic Tate module of an abelian scheme of characteristic p has less rank than its characteristic 0 counterpart. The self-duality at p of L has to be imposed to overcome this point (see the argument just above Remark 7.4). Also we need the density of the derived group $G_1(\mathbb{Q})$ in $G_1(\mathbb{A}^{(\infty)})$ (the strong approximation theorem) in order to know that geometrically irreducible components of $Sh_K^{(\Sigma)}$ are indexed by the class group $Cl(K)$: $\pi_0(Sh^{(\Sigma)}_{K/\overline{\mathbb{Q}}}) \cong Cl(K)$. For Shimura varieties of a more general pair (G, X), these points have to be studied carefully (see Remark 7.4).

If $p \in \Sigma$ is unramified in $F/\mathbb{Q}$ (and $K^{(\Sigma)}$ is sufficiently small), $Sh_K^{(\Sigma)}$ is smooth over $\mathbb{Z}_{(p)}$ by the infinitesimal criterion of smoothness (e.g., [NMD] Proposition 2.2.6); that is, we can show that any characteristic p test object lifts to characteristic 0 infinitesimally. To explain this, let R be a $\mathbb{Z}_{(p)}$-algebra with a nilpotent ideal $I \subset R$ containing a power of p. Put $R_0 = R/I$. We want to show the existence of a lifting of a test object $(A_0, \overline{\lambda}_0, \overline{\eta}_0^{(\Sigma)})_{/R_0}$ to R. The abelian variety A_0 lifts to an abelian scheme $A_{/R}$ (with $A \otimes_R R_0 \cong A_0$) by the deformation theory of Grothendieck–Mumford (cf. [CBT] V.1.6, [GIT] Section 6.3, [DAV] I.3, and also Theorem 8.8 in the text and the remark after the theorem). Since the degree of polarization is prime to p (here we use the fact that we can choose a representative $\mathfrak{c}$ prime to p in a given class in $Cl(K)$), λ also lifts because we may assume that $\lambda_0 : A_0 \to {}^tA_0$ is étale (and hence ${}^tA = A/C$ for an étale subgroup $C \subset A$ lifting $\mathrm{Ker}(\lambda_0)$; see [ECH] I.3.12). As for the level structure $\eta_0^{(\Sigma)}$, it is prime to p and hence étale over R_0. Then it extends uniquely to a level structure $\eta^{(\Sigma)} : V_{\mathbb{A}}^{(\Sigma)} \cong V^{(\Sigma)}(A)$ over R. By the deformation theory of Barsotti–Tate groups (see [CBT] V.1.6), using (det), we can find a deformation $A_{/R}$ of A_{0/R_0} with an embedding $O \hookrightarrow \mathrm{End}(A_{/R})$ compatible with $O \hookrightarrow \mathrm{End}(A_{0/R_0})$ (see Theorem 8.8 and the remark after the theorem).

When Σ consists of a single prime p, we write $\mathcal{E}_K^{(\Sigma)}$ as $\mathcal{E}_K^{(p)}$, $Sh_K^{(\Sigma)}$ as $Sh_K^{(p)}$, and so on.

4.2.2 Global Reciprocity Law

Traditionally a reciprocity law in class field theory is an explicit description of a field automorphism group (see examples in Chapter 2). The global reciprocity law of a Shimura variety is such a description of automorphisms of a geometrically connected component of the Shimura variety as a scheme over $\mathbb{Q}$. An interesting fact (particular to Shimura varieties) is that all generic automorphisms (of the function field) extend to global automorphisms (see [Mt]); so, in such a case, the description of scheme automorphisms and field automorphisms of the function field are equivalent. Here we describe Shimura's global reciprocity law for the Hilbert modular variety.

Shimura's proof of the reciprocity law is based on the analysis of the topological fundamental group of the variety as a subgroup of the analytic automorphism group of the bounded symmetric domain; in our setting here, the domain is the product $\mathfrak{Z} = \mathfrak{H}^I$, and the analytic automorphism group is given by $\mathrm{Aut}(\mathfrak{Z}) = \mathfrak{S}_I \ltimes PSL_2(\mathbb{R})^I$, where $\mathfrak{S}_I$ is the symmetric group of I permuting the components $\mathfrak{H}$ of $\mathfrak{Z}$. See [IAT] 6.6 for a typical example of his proof of Theorem 2.43 in the elliptic modular case, which can be extended to more general Shimura varieties (see [Sh5] II and [MiS]). Here we give a (partially) new proof (once suggested by Ching-Li Chai; see [H03b]) based on the analysis of the algebraic fundamental group of the variety and the following theorem of Faltings and Zarhin.

Theorem 4.13 (Faltings and Zarhin) *Let κ be a field finitely generated over the prime subfield of κ. Fix a separable closure $\overline{\kappa}$ of κ. Let $A_{/\kappa}$ and $A'_{/\kappa}$ be abelian varieties defined over κ, and write $T_\ell(A)$ and $T_\ell(A')$ for the ℓ-adic Tate modules of $A_{/\overline{\kappa}}$ and $A'_{/\overline{\kappa}}$ as $\mathrm{Gal}(\overline{\kappa}/\kappa)$-modules, where ℓ is a prime invertible in κ. Then the natural injection induces the following isomorphism,*

$$\mathrm{Hom}_{\mathrm{alg.gp}}(A_{/\kappa}, A'_{/\kappa}) \otimes_{\mathbb{Z}} \mathbb{Z}_\ell \cong \mathrm{Hom}_{\mathrm{Gal}(\overline{\kappa}/\kappa)}(T_\ell(A), T_\ell(A')),$$

and $T_\ell(A) \otimes_{\mathbb{Z}} \mathbb{Q}$ is a semi-simple module over $\mathrm{Gal}(\overline{\kappa}/\kappa)$.

When κ is a finite field, this is a classical result of Tate ([T1]), and Tate conjectured the above statement to be true; indeed, one can formulate the conjecture more generally for motives over a field of finite type using ℓ-adic étale cohomology in place of the Tate modules (see [T2]). The case where the base field has odd characteristic $p > 0$ is proved by Zarhin ([Z] and [Z1]) (and later the odd prime condition for p is removed; see [DAV] V.4.7). The characteristic 0 case is due to Faltings (see [RTP] VI.3). The proof of this theorem is not given here, since it is beyond the scope of this book.

Take an algebraically closed field $\overline{k}$ of characteristic 0. An irreducible component of $Sh(G,X) \times_{\mathbb{Q}} \overline{k}$ is called a *geometrically irreducible component of* $Sh(G,X)$. Fix an embedding $\overline{\mathbb{Q}} \hookrightarrow \overline{k}$. For each open compact subgroup K, $Sh_K(G,X) \times_{\mathbb{Q}} \overline{\mathbb{Q}}$ has finitely many irreducible components, and irreducibility is intact after extension of the scalar to $\overline{k}$ over $\overline{\mathbb{Q}}$. Thus the notion of a geometrically irreducible component of $Sh(G,X)$ is independent of the choice of $\overline{k}$.

Let us make explicit some geometrically irreducible components of the Shimura variety $Sh(G,X)$. The functor $\mathcal{E}'_{\Gamma(N)} : \mathbb{Q}\text{-}SCH \to SETS$ given by $\mathcal{E}'_{\Gamma(N)}(\overline{k}) = [(A, \lambda, \phi'_N)_{/\overline{k}}]$ with $\phi'_N : \mathfrak{d}^{-1}/N\mathfrak{d}^{-1} \times O/NO \cong A[N]$ and an O-polarization λ is representable over $\mathbb{Q}$ by $\mathfrak{M}'(O, \Gamma(N))_{/\mathbb{Q}}$ if $N \geq 3$. Let $\mathfrak{M}' = \varprojlim_N \mathfrak{M}'(O, \Gamma(N))$. The functorial morphism $\mathcal{E}'_{\Gamma(N)}(\overline{k}) \ni (A, \lambda, \phi'_N) \mapsto \langle \phi'_N(1,0), \phi'_N(0,1) \rangle_N \in \mu_N(\overline{k})$ (after taking the limit with respect to N) induces a morphism $\iota : \mathfrak{M}' \to Spec(\mathbb{Q}^{cyc})$ for $\mathbb{Q}^{cyc} = \bigcup_N \mathbb{Q}[\mu_N]$ making the following diagram commutative.

$$\begin{array}{ccc} \mathfrak{M}' & \xrightarrow{\subset} & Sh(G,X) \\ {\scriptstyle\iota}\downarrow & & \downarrow{\scriptstyle\pi} \\ Spec(\mathbb{Q}^{cyc}) & \longrightarrow & Spec(\mathbb{Q})\,. \end{array}$$

Giving ϕ'_N becomes equivalent to having a level $\Gamma(N)$-structure over $\mathbb{Q}[\zeta_N]$ under the choice of a primitive root $\zeta_N = \langle \phi'_N(1,0), \phi'_N(0,1)\rangle_N$ (which gives an identification: $\mu_N \otimes \mathfrak{d}^{-1}$ with $\mathfrak{d}^{-1}/N\mathfrak{d}^{-1}$). Thus we find that

$$\mathfrak{M}'(O,\Gamma(N)) \times_{\mathbb{Q}} \mathbb{Q}[\mu_N] \cong \bigsqcup_{\zeta_N} \mathfrak{M}(O,\Gamma(N)) \times_{\mathbb{Q}} \mathbb{Q}[\mu_N], \tag{4.27}$$

and hence $\mathfrak{M}'(O,\Gamma(N))_{/\mathbb{Q}}$ is isomorphic to $\mathfrak{M}(O,\Gamma(N)) \times_{\mathbb{Q}} \mathbb{Q}[\mu_N]$ regarded as defined over $\mathbb{Q}$. Note that $\mathfrak{M}'(O,\Gamma(N)) \hookrightarrow Sh_{S(N)}(G,X)$ by our construction of the Shimura variety. The variety $\mathfrak{M}(O,\Gamma(N))_{/\mathbb{Q}}$ is geometrically irreducible, because of the connectedness over $\mathbb{C}$: $\mathfrak{M}(O,\Gamma(N))(\mathbb{C}) = \Gamma(O,N)\backslash\mathfrak{Z}$. Since $\{S(N)\}_N$ gives a system of open neighborhoods of the identity of $G(\mathbb{A}^{(\infty)})$,

$$\mathfrak{M}_{/\mathbb{Q}^{cyc}} = \varprojlim_N \mathfrak{M}(O,\Gamma(N)) \times_{\mathbb{Q}} \mathbb{Q}[\zeta_N]$$

for a compatible system of Nth roots of unity so that $\zeta_m = \zeta_n^{n/m}$ (if $m|n$) gives one geometrically irreducible component, which is the image of $X^+ \times 1$ in the expression (4.21) over $\mathbb{C}$. This component $\mathfrak{M}$ is called the *neutral component* of the Shimura variety ([D2] Définition 2.1.5).

Let $\widetilde{c} = \left(\begin{smallmatrix} c & 0 \\ 0 & 1 \end{smallmatrix}\right)$ for $c \in F^{\times}_{\mathbb{A}^{(\infty)}}$. Since $\Gamma(\mathfrak{c},N) = \left(\widetilde{c}S(N)\widetilde{c}^{-1}\right) \cap G_1(\mathbb{Q})$ if $c\widehat{O} \cap F = \mathfrak{c}$, we find that

$$\mathfrak{M}_{\mathfrak{c}/\mathbb{Q}^{cyc}} = \varprojlim_N \mathfrak{M}(\mathfrak{c},\Gamma(N)) \times_{\mathbb{Q}} \mathbb{Q}[\zeta_N] \tag{4.28}$$

is the image of $\mathfrak{M}$ under the right action of $\widetilde{c}$; so, $\mathfrak{M}_{\mathfrak{c}}$ gives another geometrically irreducible component of $Sh(G,X)_{/\mathbb{Q}^{cyc}}$. In other words, $\mathbb{Q}^{cyc}$ is the field of definition of the component $\mathfrak{M}_{\mathfrak{c}}$ in the sense that $\mathbb{Q}^{cyc}$ is the algebraic closure of $\mathbb{Q}$ in the function field of $\mathfrak{M}_{\mathfrak{c}}$ over $\mathbb{Q}$.

By the description (4.21) of complex points of the Shimura variety, the group $G(\mathbb{A}^{(\infty)})$ acts transitively on $\pi_0(Sh(G,X)_{/\overline{\mathbb{Q}}})$. Thus, up to conjugation by $G(\mathbb{A}^{(\infty)})$, the automorphism group of a geometrically irreducible component is independent of the choice of the component.

We now start writing down explicitly the group that gives an open subgroup of $\mathrm{Aut}(\mathfrak{M}_{/\mathbb{Q}})$ with finite index. As before we let $g \in G(\mathbb{A}^{(\infty)})$ act on $Sh(G,X)_{/\mathbb{Q}}$ by

$$(A,\overline{\lambda},\eta) \mapsto (A,\overline{\lambda},\eta \circ g),$$

which gives a right action of $G(\mathbb{A}^{(\infty)})$ on $Sh(G,X)$. We write $\tau(g)$ for the automorphism of $Sh(G,X)_{/\mathbb{Q}}$ induced by $g \in G(\mathbb{A}^{(\infty)})$. For each point $x = (A,\overline{\lambda},\eta)$ of the Shimura variety $Sh(G,X)$, we can associate a lattice $\widehat{L} = \eta^{-1}(\mathcal{T}(A)) \subset F^2_{\mathbb{A}^{(\infty)}}$, and a choice of a base $w = (w_1,w_2)$ of this lattice

gives the identification $\eta : \widehat{O}^2 \cong \mathcal{T}(A)$ by $(a,b) \mapsto a\eta(w_1) + b\eta(w_2)$. This is actually how we related the two functors $\mathcal{E}_K^{(\emptyset)}$ and $\mathcal{E}'_K$. In other words, the automorphism $x \mapsto \tau(g)(x)$ is interpreted into the base change: ${}^t w \mapsto g^{-1} \cdot {}^t w$ because $(\eta \circ g)^{-1}(\mathcal{T}(A)) = g^{-1}\eta^{-1}(\mathcal{T}(A)) = g^{-1}\widehat{L}$. Over $\mathbb{C}$, test object (A, λ, ω) is classified by lattices $\mathcal{L} \subset \mathbb{C}^I$ and its base w. Thus the action of $\tau(\alpha)$ for $\alpha \in G(\mathbb{Q})_+ \cap G(\widehat{\mathbb{Z}})$ induces $z \mapsto \alpha^{-1}(z)$ (see also (4.2)). Since in Theorem 2.43 we normalized the action using the action on lattices, it is given by $\tau(g)^{-1}$ and hence $GL_2(\mathbb{A})$ acts on the elliptic modular function field $\mathfrak{K}$ on the right. In the corresponding theorem, Theorem 4.14 in the Hilbert modular case, the action is on the left given by the pullback action of $\tau(g)$.

Here is the description by Shimura of a group that gives almost all automorphisms of the neutral component $\mathfrak{M}_{/\mathbb{Q}}$:

$$\mathcal{G}(G,X) = \left\{ g \in G(\mathbb{A}) \middle| \det(g) \in \mathbb{A}^\times \overline{F^\times F_{\infty+}^\times} \right\},$$

and define $\overline{\mathcal{E}} = \overline{\mathcal{E}}(G,X) = \mathcal{G}(G,X)/\overline{Z(\mathbb{Q})G(\mathbb{R})_+}$ (see [Sh5] II, [Sh6], and [AAF] Section 8). Here $F_{\infty+}^\times$ is the subgroup of totally positive elements in $F_\infty = F \otimes_{\mathbb{Q}} \mathbb{R}$. This group $\overline{\mathcal{E}}(G,X)$ gives a subgroup of finite index in the automorphism group in question.

For a proper subset Σ of rational primes, we define $G(\mathbb{Z}_\Sigma) = \prod_{p \in \Sigma} G(\mathbb{Z}_p)$, which is a closed subgroup of $G(\mathbb{A}^{(\infty)})$. We have the quotient variety

$$Sh^{(\Sigma)}(G,X) = Sh(G,X)/G(\mathbb{Z}_\Sigma). \tag{Σ}$$

We write $\mathfrak{M}_{/\mathbb{Q}}^{(\Sigma)}$ for the image of $\mathfrak{M}$ in $Sh^{(\Sigma)}(G,X)$. We define the localization of $\mathbb{Z}$ at Σ by $\mathbb{Z}_{(\Sigma)} = \mathbb{Q} \cap \mathbb{Z}_\Sigma$ inside $\mathbb{A}^{(\infty)}$ for $\mathbb{Z}_\Sigma = \prod_{p \in \Sigma} \mathbb{Z}_p$. As a convention, we write $\mathbb{Z}_{(\emptyset)} = \mathbb{Q}$ and $\mathbb{Z}_\emptyset = \mathbb{A}^{(\infty)}$. We define

$$\mathcal{G}^{(\Sigma)}(G,X) = \left\{ g \in G(\mathbb{A}^{(\Sigma)}) \middle| \det(g) \in {\mathbb{A}^{(\Sigma)}}^\times \overline{O_{(\Sigma)}^\times F_{\infty+}^\times} \right\},$$

where ${\mathbb{A}^{(\Sigma)}}^\times = \{x \in \mathbb{A}^\times | x_p = 1 \text{ if } p \in \Sigma\}$ and $O_{(\Sigma)} = O \otimes_{\mathbb{Z}} \mathbb{Z}_{(\Sigma)} \subset F$. We put

$$\overline{\mathcal{E}}^{(\Sigma)} = \overline{\mathcal{E}}^{(\Sigma)}(G,X) = \mathcal{G}^{(\Sigma)}(G,X)/\overline{Z(\mathbb{Z}_{(\Sigma)})G(\mathbb{R})_+},$$

which gives almost all automorphisms of the component $\mathfrak{M}_{/\mathbb{Q}}^{(\Sigma)}$ of $Sh^{(\Sigma)}$.

We recall here another definition of $\overline{\mathcal{E}}$ due to Deligne [D2] (see also [MiS]), although in the present Hilbert modular case, the definition given here looks (and indeed is) more complicated than the one already given. However, this new definition works well for any algebraic group G that admits Shimura varieties. So we introduce general notation for reductive group $G_{/\mathbb{Q}}$, forgetting for the moment about our specific group $\mathrm{Res}_{F/\mathbb{Q}}GL(2)$. Recall the identity-connected component $G(\mathbb{R})^+$ of the Lie group $G(\mathbb{R})$ and $G(\mathbb{Q})^+ = G(\mathbb{Q}) \cap G(\mathbb{R})^+$. Write $G^{ad} = PG$ for the adjoint group of G (thus

G^{ad} is the algebraic group quotient of G modulo the center Z, but $G^{ad}(k)$ may not equal $G(k)/Z(k)$ depending on the field k). For the quotient map $ad : G \to G^{ad}$, we define $G(\mathbb{Q})_+ = ad^{-1}(G^{ad}(\mathbb{Q})^+)$. Often $G(\mathbb{Q})^+ = G(\mathbb{Q})_+$, but there are examples where $G(\mathbb{Q})_+ \supsetneq G(\mathbb{Q})^+$ (like some unitary groups). Here we use the symbol G^{ad} in place of PG in order to facilitate an easy reference to the papers quoted above ([D2] and [MiS]).

We write $\overline{Z(\mathbb{Q})}$ for the topological closure of the center $Z(\mathbb{Q})$ of $G(\mathbb{Q})$ in $G(\mathbb{A}^{(\infty)})$. Let the adjoint group $G^{ad}(\mathbb{Q})$ act on $G(\mathbb{A}^{(\infty)})/\overline{Z(\mathbb{Q})}$ by conjugation and consider the semi-direct product

$$\frac{G(\mathbb{A}^{(\infty)})}{\overline{Z(\mathbb{Q})}} \rtimes G^{ad}(\mathbb{Q})^+ = \left\{ (g,h) \middle| g \in G(\mathbb{A}^{(\infty)})/\overline{Z(\mathbb{Q})},\ h \in G^{ad}(\mathbb{Q})^+ \right\},$$

on which the product law is given by $(g,h)(g',h') = (g \cdot ad(h)(g'), hh')$, where $ad(h)(g') = hg'h^{-1}$. We now bring the group $G(\mathbb{Q})_+$ into the semi-direct product $\frac{G(\mathbb{A}^{(\infty)})}{\overline{Z(\mathbb{Q})}} \rtimes G^{ad}(\mathbb{Q})^+$ by $\gamma \mapsto (\gamma^{-1}, ad(\gamma))$. Then the image is a normal subgroup of $\frac{G(\mathbb{A}^{(\infty)})}{\overline{Z(\mathbb{Q})}} \rtimes G^{ad}(\mathbb{Q})^+$. We define the amalgamated product $\frac{G(\mathbb{A}^{(\infty)})}{\overline{Z(\mathbb{Q})}} *_{G(\mathbb{Q})_+} G^{ad}(\mathbb{Q})^+$ to be the quotient of $\frac{G(\mathbb{A}^{(\infty)})}{\overline{Z(\mathbb{Q})}} \rtimes G^{ad}(\mathbb{Q})^+$ by the image of $G(\mathbb{Q})_+$. We write $[g, ad(\gamma)]$ for the class in $\frac{G(\mathbb{A}^{(\infty)})}{\overline{Z(\mathbb{Q})}} *_{G(\mathbb{Q})_+} G^{ad}(\mathbb{Q})^+$ represented by $(g, ad(\gamma)) \in \frac{G(\mathbb{A}^{(\infty)})}{\overline{Z(\mathbb{Q})}} \rtimes G^{ad}(\mathbb{Q})^+$. Since

$$\frac{G(\mathbb{A}^{(\infty)})}{\overline{Z(\mathbb{Q})}} *_{G(\mathbb{Q})_+} G^{ad}(\mathbb{Q})^+ \ni [g, ad(\gamma)] \mapsto \det(g) \in F_{\mathbb{A}}^{\times}/\overline{F^{\times}F_{\infty+}^{\times}}$$

is well-defined, we define a new group $\overline{\mathcal{E}}'(G,X)$ by the inverse image of the closed subgroup $\mathbb{A}^{\times}\overline{F^{\times}F_{\infty+}^{\times}}/\overline{F^{\times}F_{\infty+}^{\times}}$ of $F_{\mathbb{A}}^{\times}/\overline{F^{\times}F_{\infty+}^{\times}}$ in the amalgamated product $\frac{G(\mathbb{A}^{(\infty)})}{\overline{Z(\mathbb{Q})}} *_{G(\mathbb{Q})_+} G^{ad}(\mathbb{Q})^+$.

We may axiomatize the above construction as follows (cf. [D2] 2.0.1). Take three groups $\mathcal{G}$, Γ, and Δ. We assume to have a homomorphism $\Gamma \ni \gamma \mapsto \overline{\gamma}$ onto a subgroup $\overline{\Gamma}$ of $\mathcal{G}$ and to have the following commutative diagram of group homomorphisms.

$$\begin{array}{ccc} \overline{\Gamma} & \xrightarrow{\varphi} & \Delta \\ \cap\downarrow & & \downarrow r \\ \mathcal{G} & \xrightarrow[ad]{} & \mathrm{Aut}(\mathcal{G}). \end{array} \tag{4.29}$$

Here $\mathrm{Aut}(\mathcal{G})$ is the automorphism group of the group $\mathcal{G}$ and $ad(g)(x) = gxg^{-1}$ for $g \in \mathcal{G}$. We often write $\varphi(\gamma)$ for $\varphi(\overline{\gamma})$ and assume that $r(\delta)$ ($\delta \in \Delta$) preserves $\overline{\Gamma}$ as a whole. We suppose the following two compatibility conditions,

(a) $r(\varphi(\gamma)) = ad(\overline{\gamma})$ for all $\gamma \in \Gamma$ (commutativity of (4.29));

(b) $\varphi(r(\delta)(\overline{\gamma})) = ad(\delta)(\varphi(\gamma))$ for all $\delta \in \Delta$ and $\gamma \in \Gamma$, where $ad(x)(y) = xyx^{-1}$.

We consider the semi-direct product: $\mathcal{G} \rtimes \Delta$ whose multiplication law is given by $(g, \delta)(h, \epsilon) = (g \cdot (r(\delta)(h)), \delta\epsilon)$, and we have $(g, \delta)^{-1} = (r(\delta^{-1})(g^{-1}), \delta^{-1})$. By computation, we have

$$(g, \delta)(\overline{\gamma}^{-1}, \varphi(\gamma))(g, \delta)^{-1} = (g \cdot r(\delta)(\overline{\gamma}^{-1}) r(\delta \cdot \varphi(\gamma))(r(\delta^{-1})(g^{-1})), ad(\delta)(\varphi(\gamma))).$$

Then again by computation,

$$\begin{aligned} & g \cdot r(\delta)(\overline{\gamma}^{-1}) r(\delta \cdot \varphi(\gamma))(r(\delta^{-1})(g^{-1})) \\ & \quad = g \cdot r(\delta)(\overline{\gamma}^{-1}) r(\delta)(\overline{\gamma}(r(\delta^{-1})(g^{-1}))\overline{\gamma}^{-1}) = g \cdot r(\delta)((r(\delta^{-1})(g^{-1}))\overline{\gamma}^{-1}) \\ & \qquad = g \cdot r(\delta)(r(\delta^{-1})(g^{-1}))(r(\delta)(\overline{\gamma}^{-1})) = r(\delta)(\overline{\gamma})^{-1}. \end{aligned}$$

This shows that $\widetilde{\Gamma} = \{(\overline{\gamma}^{-1}, \varphi(\gamma)) | \gamma \in \Gamma\}$ is a normal subgroup of the semi-direct product $\mathcal{G} \rtimes \Delta$. We then define

$$\mathcal{G} *_\Gamma \Delta = (\mathcal{G} \rtimes \Delta)/\widetilde{\Gamma}. \tag{4.30}$$

We have the following commutative diagram with exact rows.

$$\begin{array}{ccccccc} \mathrm{Ker}(\varphi) & \xrightarrow{\subset} & \overline{\Gamma} & \xrightarrow{\varphi} & \Delta & \xrightarrow{\text{onto}} & \mathrm{Coker}(\varphi) \\ \| \downarrow & & \cap \downarrow & & \cap \downarrow{\scriptstyle \delta \mapsto [1,\delta]} & & \downarrow \| \\ \mathrm{Ker}(\varphi) & \xrightarrow[\subset]{} & \mathcal{G} & \xrightarrow[g \mapsto [g,1]]{} & \mathcal{G} *_\Gamma \Delta & \xrightarrow[\text{onto}]{} & \mathrm{Coker}(\varphi). \end{array}$$

Then by the suitably applied snake lemma, we get a canonical isomorphism

$$\overline{\Gamma}\backslash\mathcal{G} \cong \Delta\backslash(\mathcal{G} *_\Gamma \Delta). \tag{4.31}$$

By this isomorphism, the amalgamated product $\mathcal{G} *_\Gamma \Delta$ acts on $\overline{\Gamma}\backslash\mathcal{G}$, and the action of $[g, \delta] \in \mathcal{G} *_\Gamma \Delta$ on the class $[x]$ in $\overline{\Gamma}\backslash\mathcal{G}$ (which is sent to $[x, 1] \in \Delta\backslash(\mathcal{G} *_\Gamma \Delta)$) is given by

$$[x] \cdot [g, \delta] = [x, 1][g, \delta] = [xg, \delta] = [1, \delta][r(\delta)^{-1}(xg), 1] = [r(\delta)^{-1}(xg), 1]. \tag{4.32}$$

Suppose that $\mathcal{G}$ is a topological group with continuous action r of Δ on $\mathcal{G}$ under the discrete topology on Δ and that $\overline{\Gamma}$ is a closed subgroup of $\mathcal{G}$. Giving the group $\mathcal{G} *_\Gamma \Delta$ the quotient topology of the product topology on $G \rtimes \Delta$, $\mathcal{G} *_\Gamma \Delta$ becomes a topological group, and $\mathcal{G}/\mathrm{Ker}(\varphi) \subset \mathcal{G} *_\Gamma \Delta$ is an open subgroup. Then (4.31) is a homeomorphism, which is left to the reader to verify.

Let us now return to $G = \mathrm{Res}_{O/\mathbb{Z}} GL(2)$ and $G_1 = \mathrm{Res}_{O/\mathbb{Z}} SL(2)$. Since

$$Z(\mathbb{A}^{(\infty)}) \cap G(\mathbb{Q}) = \overline{Z(\mathbb{Q})} \cap G(\mathbb{Q}) = Z(\mathbb{Q}),$$

we see that $G^{ad}(\mathbb{Q}) = G(\mathbb{Q})/Z(\mathbb{Q})$ is a discrete subgroup of $\frac{G(\mathbb{A}^{(\infty)})}{\overline{Z(\mathbb{Q})}}$. Taking

$$(\mathcal{G}, \Gamma, \Delta) = (\frac{G(\mathbb{A}^{(\infty)})}{\overline{Z(\mathbb{Q})}}, G(\mathbb{Q})_+, G^{ad}(\mathbb{Q})^+)$$

with φ given by $\gamma \mapsto ad(\gamma)$ and $\gamma \mapsto \overline{\gamma}$ given by $\gamma \mapsto (\gamma \mod \overline{Z(\mathbb{Q})})$, we find

$$\frac{G(\mathbb{A}^{(\infty)})}{\overline{Z(\mathbb{Q})}} *_{G(\mathbb{Q})_+} G^{ad}(\mathbb{Q})^+ = \frac{G(\mathbb{A}^{(\infty)})}{\overline{Z(\mathbb{Q})}} *_{G^{ad}(\mathbb{Q})^+} G^{ad}(\mathbb{Q})^+.$$

From this it is obvious that the above group is $\frac{G(\mathbb{A}^{(\infty)})}{\overline{Z(\mathbb{Q})}}$, and $\overline{\mathcal{E}}(G,X)$ coincides with $\overline{\mathcal{E}}'(G,X)$. Thus the two definitions of $\overline{\mathcal{E}}$ are consistent. We record

$$\mathcal{G}(G,X)/\overline{Z(\mathbb{Q})}G(\mathbb{R})_+ = \overline{\mathcal{E}}(G,X) \subset \frac{G(\mathbb{A}^{(\infty)})}{\overline{Z(\mathbb{Q})}} *_{G(\mathbb{Q})_+} G^{ad}(\mathbb{Q})^+ \qquad (4.33)$$

by $g \leftrightarrow [g,1]$.

Take the geometrically irreducible component $\mathfrak{M}_{/\mathbb{Q}^{cyc}} \subset Sh(G,X)_{/\mathbb{Q}^{cyc}}$ containing the image of $X^+ \times 1$, and define $\mathrm{Aut}(\mathfrak{M}_{/\mathbb{Q}})$ by the scheme automorphism group. The following fact has been shown in [Sh5] II 6.5 and [Mt] Theorem 2 (see also [MiS] 4.6 and 4.13).

Theorem 4.14 *Let Σ be a proper subset of rational primes (Σ could be empty). The stabilizer in $G(\mathbb{A}^{(\Sigma\infty)})/\overline{Z(\mathbb{Z}_{(\Sigma)})}$ of each geometrically irreducible component of $Sh^{(\Sigma)}(G,X)_{/\mathbb{Q}}$ is given by $\overline{\mathcal{E}}^{(\Sigma)}(G,X)$. When $\Sigma = \emptyset$, the right action of $(g, ad(\gamma)) \in \overline{\mathcal{E}}(G,X)$ ($\gamma \in G(\mathbb{Q})_+$) on $[x,g']$ is given by $[x,g'] \mapsto [\gamma^{-1}(x), (g'g)^{ad(\gamma)}]$, where $(g'g)^{ad(\gamma)} = \gamma^{-1}(g'g)\gamma$. Moreover for the function field $\mathfrak{K}^{(\Sigma)}$ of $\mathfrak{M}^{(\Sigma)}_{/\mathbb{Q}}$, $\mathrm{Aut}(\mathfrak{M}^{(\Sigma)}_{/\mathbb{Q}}) = \mathrm{Aut}(\mathfrak{K}^{(\Sigma)})$, and we have a canonically splitting exact sequence of topological groups:*

$$1 \to \overline{\mathcal{E}}^{(\Sigma)}(G,X) \xrightarrow{\tau} \mathrm{Aut}(\mathfrak{K}^{(\Sigma)}) \to \mathrm{Aut}(F) \to 1,$$

where $\mathrm{Aut}(F)$ is the group of field automorphisms of F. The splitting of the above projection $\mathrm{Aut}(\mathfrak{K}^{(\Sigma)}) \twoheadrightarrow \mathrm{Aut}(F)$ is induced by the action of $\sigma \in \mathrm{Aut}(F)$ on test objects given by $(A, \overline{\lambda}, i : O \hookrightarrow \mathrm{End}(X), \eta) \mapsto (A, \overline{\lambda}, i \circ \sigma, \eta \circ \sigma)$.

The neutral component $\mathfrak{M}^{(\emptyset)}$ in the theorem is the tower of canonical models in the sense of Shimura; so, the above result includes an interpretation in Deligne's language of a result of Shimura in [Sh5] II 6.5. When we regard $g \in \overline{\mathcal{E}}(G,X)$ as an automorphism of $\mathcal{O}_{Sh}$ or $Sh(G,X)_{/\mathbb{Q}}$, we write it as $\tau(g)$.

The topology of the Galois groups is as introduced in Section 2.3; in particular, $\mathrm{Aut}(F)$ has the discrete topology. Note that, a priori, $\mathrm{Aut}(\mathfrak{M}^{(\Sigma)}_{/\mathbb{Q}})$ is a subgroup of $\mathrm{Aut}(\mathfrak{K}^{(\Sigma)})$ for the function field $\mathfrak{K}^{(\Sigma)}$ of $\mathfrak{M}^{(\Sigma)}$, and we know, a posteriori, the equality of the two. In any case, the topology of $\mathrm{Aut}(\mathfrak{M}^{(\Sigma)}_{/\mathbb{Q}})$ is induced by that of $\mathrm{Aut}(\mathfrak{K}^{(\Sigma)})$ given in Section 2.3.

The homomorphism τ in the theorem is normalized with respect to $G(\mathbb{A}^{(\infty)})$, and the present $\tau(x)$ is actually $\tau(x)^{-1}$ in Theorem 2.43, because the action on the level structure $\eta \mapsto \eta \circ \alpha^{(\infty)}$ corresponds to $z \mapsto \alpha^{-1}(z)$ by (4.2). This comes from the fact that we took the column vector space $V = F^2$ with left $G(\mathbb{Q})$-action in order to define the level structures η. Because of this, we embedded $G(\mathbb{Q})_+$ into $G(\mathbb{A}^{(\infty)}) \rtimes G^{ad}(\mathbb{Q})$ by $\alpha \mapsto (\alpha^{-1}, ad(\alpha))$.

Over $\mathbb{C}$, by (4.21), we have $[z, g] = [\gamma(z), \gamma g]$ ($\Leftrightarrow [\gamma^{-1}(z), g] = [z, \gamma g]$) for $(z, g) \in \mathfrak{Z} \times G(\mathbb{A}^{(\infty)})$ and $\gamma \in G(\mathbb{Q})_+$, taking the expression

$$Sh(G, X)(\mathbb{C}) = G(\mathbb{Q})_+ \backslash \left(X^+ \times G(\mathbb{A}^{(\infty)}) \right) / \overline{Z(\mathbb{Q})}$$

and noting $X^+ = \mathfrak{Z}$. Each $[z, g]$ carries the triple $(A_z, \lambda_z, \eta_z \circ g)$, where $A_z(\mathbb{C}) = \mathbb{C}^I/(O^* + Oz)$ and $\eta_z \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) = bz - a$ identifying $\mathcal{T}(A_z) = \widehat{O}^* + \widehat{O}z$. To see this, we note an easy fact that the map: $[z, g] \mapsto (A_z, \lambda_z, \eta_z \circ g)$ sends $\left(\mathfrak{Z} \times G(\mathbb{A}^{(\infty)})\right)$ surjectively onto $\mathcal{E}^{(\emptyset)}(\mathbb{C})$. Thus we need to check

$$(A_{\gamma^{-1}(z)}, \overline{\lambda}_{\gamma^{-1}(z)}, \eta_{\gamma^{-1}(z)} \circ g) \cong (A_z, \overline{\lambda}_z, \eta_z \circ \gamma^{(\infty)} g) \text{ in } \mathcal{A}_F^{(\emptyset)} \text{ for } \gamma \in G(\mathbb{Q})_+$$

which is equivalent to $[\gamma^{-1}(z), g] = [z, \gamma^{(\infty)} g]$. This is exactly (4.2), because $\alpha_\gamma \circ \eta_{\gamma^{-1}(z)} = \eta_z \circ \gamma^{(\infty)}$ for the isogeny $\alpha_\gamma : A_{\gamma^{-1}(z)} \to A_z$ given via the multiplication by $(-cz + a)$ on $\mathbb{C}^I$ (writing $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$). This gives a sketch of the verification of $Sh(G, X)(\mathbb{C}) \cong \mathcal{E}_1^{(\emptyset)}(\mathbb{C})$ which actually has to be done for each open compact subgroup $K \subset G(\mathbb{A}^{(\infty)})$ (then taking the limit of isomorphisms with respect to K).

Proof of Theorem 4.14. We now give a proof of the theorem, which also supplies a proof of the elliptic reciprocity law formulated in Theorem 2.43. By the description of the complex points $Sh(\mathbb{C})$ ((4.21)) combined with the identity $Sh(G, X)(\mathbb{C}) \cong \mathcal{E}_1^{(\emptyset)}(\mathbb{C})$, $G(\mathbb{A})$ acts transitively on the geometrically irreducible components $\pi_0(Sh(G, X))$. Since $\overline{\mathcal{E}}^{(\Sigma)}(G, X)$ is normalized by $G(\mathbb{A}^{(\Sigma\infty)})$, we only need to prove the theorem for the neutral component $\mathfrak{M}^{(\Sigma)}$. The beginning of our proof follows the argument given in [IAT] Chapter 6, and then we follow a different path (given in [H03b]).

Our proof of Theorem 4.14 goes through several steps.

Step 1: We first remark that the description of the action of $\overline{\mathcal{E}}(G, X)$ (defined by Deligne's amalgamated product) on the variety follows from (4.32), if the functorial action $(A, \overline{\lambda}, \eta) \mapsto (A, \overline{\lambda}, \eta \circ g)$ of $g \in G(\mathbb{A}^{(\infty)})$ induces $x \mapsto xg$ on the complex points of the Shimura variety. This point has already been verified (see (4.2) and the above argument).

Step 2: $\overline{\mathcal{E}}(G, X)$ gives the stabilizer of the geometrically irreducible component $\mathfrak{M}^{(\Sigma)}$ (and hence also of $\mathfrak{M}_{\mathfrak{c}}^{(\Sigma)}$ for all $\mathfrak{c}$) in $G(\mathbb{A}^{(\Sigma\infty)})/\overline{Z(\mathbb{Z}_{(\Sigma)})}$.

By definition, we have $Sh(G, X) = \varprojlim_S Sh_S(G, X)$ for open compact subgroups S of $G(\mathbb{A}^{(\infty)})$ and, hence, the action of $G(\mathbb{A}^{(\infty)})$ on $Sh(G, X)$ is continuous. Pick $\alpha \in Z(\mathbb{Q}) = F^\times$. Then the isogeny $\alpha : A \to A$ sends

$$(A, \overline{\lambda}, \eta) \ \text{ to } \ (A, \overline{\lambda} = \overline{\alpha^2 \lambda}, \alpha\eta = \eta \circ \alpha^{(\infty)}).$$

So, the two test objects $(A, \overline{\lambda}, \eta)$ and $(A, \overline{\lambda}, \alpha\eta = \eta\alpha^{(\infty)})$ give rise to the same point of $Sh(G, X)$. Thus $\overline{Z(\mathbb{Q})}$ in $G(\mathbb{A}^{(\infty)})$ acts trivially by continuity, and we obtain a homomorphism $G(\mathbb{A}^{(\infty)})/\overline{Z(\mathbb{Q})} \xrightarrow{\tau} \mathrm{Aut}(Sh(G, X))$. The injectivity of τ follows either from the expression (4.21) or from the fact $\mathrm{End}^{\mathbb{Q}}(\mathbb{X}) = F$ for the universal abelian scheme $\mathbb{X}_{/\mathfrak{M}^{(\Sigma)}}$ (because $\mathbb{X}$ specializes to infinitely many nonisogenous abelian varieties with complex multiplication by distinct quadratic extensions of F).

For a given lattice $L \subset V$ and $g^{(\infty)} \in G(\mathbb{A}^{(\infty)})$, we note $g^{(\infty)}\widehat{L} \wedge g^{(\infty)}\widehat{L} = \det(g^{(\infty)})(\widehat{L} \wedge \widehat{L})$; so, if $g^{(\infty)}$ preserves $\mathfrak{M}_{\mathfrak{c}}$, $\det(g^{(\infty)})\widehat{\mathfrak{c}} \cap F$ is strictly equivalent to $\widehat{\mathfrak{c}} \cap F$. Thus $\det(g^{(\infty)}) = u\alpha^{(\infty)}$ for $\alpha \in F_+^\times$ and $u \in \widehat{O}^\times$. By choosing $\gamma \in G(\mathbb{Q})_+$ so that $\det(\gamma) = \alpha^{-1}$, we find $\gamma^{(\infty)} g^{(\infty)}$ has determinant in $\widehat{O}^\times$.

Let $\mathbb{Q}^{(\Sigma)} \subset \mathbb{Q}^{cyc}$ be the maximal extension unramified at all $p \in \Sigma$. Thus it is a union of the cyclotomic field $\mathbb{Q}[\mu_N]$ for N prime to Σ. Since $\mathfrak{M}_{\mathfrak{c}}^{(\Sigma)}$ (for $\mathfrak{c}$ prime to Σ) is irreducible and also geometrically irreducible over $\mathbb{Q}^{(\Sigma)}$, writing its function field as $\mathfrak{K}^{(\Sigma)} = \mathbb{Q}^{(\Sigma)}(\mathfrak{M}_{\mathfrak{c}}^{(\Sigma)})$, any automorphism of $\mathfrak{K}^{(\Sigma)}$ over $\mathbb{Q}^{(\Sigma)}$ extends to $\mathfrak{K}_{\mathbb{C}}^{(\Sigma)} = \mathfrak{K}^{(\Sigma)} \otimes_{\mathbb{Q}^{(\Sigma)}} \mathbb{C}$, which is the function field of $\mathfrak{M}_{\mathfrak{c}/\mathbb{C}}$; thus we have the following facts.

$$\begin{aligned} &1 \to \mathrm{Aut}(\mathfrak{K}^{(\Sigma)}/\mathbb{Q}^{(\Sigma)}) \to \mathrm{Aut}(\mathfrak{K}^{(\Sigma)}) \xrightarrow{\mathrm{Res}} \mathrm{Gal}(\mathbb{Q}^{(\Sigma)}/\mathbb{Q}) \to 1 \ \text{ is exact;} \\ &\mathrm{Aut}(\mathfrak{K}^{(\Sigma)}/\mathbb{Q}^{(\Sigma)}) \cong \mathrm{Aut}(\mathfrak{K}_{\mathbb{C}}^{(\Sigma)}/\mathbb{C}). \end{aligned} \tag{4.34}$$

Since the minimal compactification $M^*(\mathfrak{c}, \Gamma(N))$ of $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ over $\mathbb{Q}$ is given by $Proj(G(\mathfrak{c}, \Gamma(N); \mathbb{Q}))$ for $G(\mathfrak{c}, \Gamma(N); \mathbb{Q}) = \bigoplus_{0 \le j} G_{jI}(\mathfrak{c}, \Gamma(N); \mathbb{Q})$, we have for N prime to Σ,

$$M^*(\mathfrak{c}, \Gamma(N)) \times_{\mathbb{Q}} \mathbb{Q}^{(\Sigma)} = Proj(G(\mathfrak{c}, \Gamma(N); \mathbb{Q}) \otimes_{\mathbb{Q}} \mathbb{Q}^{(\Sigma)}).$$

Thus we can let $\mathrm{Gal}(\mathbb{Q}^{(\Sigma)}/\mathbb{Q})$ act on $M^*(\mathfrak{c}, \Gamma(N))_{/\mathbb{Q}}$ through the $\mathbb{Q}^{(\Sigma)}$-factor of the tensor product $G(\mathfrak{c}, \Gamma(N); \mathbb{Q}) \otimes_{\mathbb{Q}} \mathbb{Q}^{(\Sigma)}$. By this "$Proj$" expression, the arithmetic action of $\mathrm{Gal}(\mathbb{Q}^{(\Sigma)}/\mathbb{Q})$ preserves the open subscheme $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$. Thus we have a canonical section s of the restriction map $\mathrm{Aut}(\mathfrak{M}_{/\mathbb{Q}}) \xrightarrow{\mathrm{Res}} \mathrm{Gal}(\mathbb{Q}^{(\Sigma)}/\mathbb{Q})$. By the q-expansion principle, the action of $s(\sigma)$ on $f(q) = \sum_\xi a(\xi) q^\xi \in G_{jI}(\mathfrak{c}, \Gamma(N); \mathbb{Q}^{(\Sigma)})$ is induced by $\sum_\xi a(\xi) q^\xi \mapsto \sum_\xi a(\xi)^\sigma q^\xi$. On the other hand, by (4.27), the action of $\widetilde{c} = \left(\begin{smallmatrix} c & 0 \\ 0 & 1 \end{smallmatrix}\right)$ for $c \in \widehat{\mathbb{Z}}^\times$, is given by $\phi'_N(x, y) \mapsto \phi'_N(cx, y)$ for ${}^t(x, y) \in \mathfrak{d}^{-1}/N\mathfrak{d}^{-1} \times O/NO$, which after identifying $\mu_N \otimes \mathfrak{d}^{-1}$ with $\mathfrak{d}^{-1}/N\mathfrak{d}^{-1}$ by choosing $\zeta_N \in \mu_N$, corresponds to the automorphism $\sigma \in \mathrm{Gal}(\mathbb{Q}[\mu_N]/\mathbb{Q})$ with $(\zeta_N^c)^\sigma = \zeta_N$ (note here the action of a scheme automorphism on the structure sheaf is the pullback action). Identifying $\widehat{\mathbb{Z}}^\times$ with $\mathrm{Gal}(\mathbb{Q}^{cyc}/\mathbb{Q})$ as in (2.2), we find $\mathrm{Gal}(\mathbb{Q}^{(\Sigma)}/\mathbb{Q}) \cong \widehat{\mathbb{Z}}^{(\Sigma)\times}$, since $\mathbb{Z}_p^\times \subset \widehat{\mathbb{Z}}^\times$ gives the inertia group at $p \in \Sigma$. Thus for a section $s_{\mathbb{A}} : \widehat{\mathbb{Z}}^{(\Sigma)\times} \hookrightarrow G(\mathbb{A}^{(\Sigma\infty)})/\overline{Z}(\mathbb{Z}_{(\Sigma)})$ given by $s_{\mathbb{A}}(c) = \widetilde{c}$, we have the following commutative diagram,

$$\begin{array}{ccc} G(\mathbb{A}^{(\Sigma\infty)})/\overline{Z(\mathbb{Z}_{(\Sigma)})} & \xrightarrow{\subset} & \mathrm{Aut}(Sh^{(\Sigma)}(G,X)) \\ {\scriptstyle s_{\mathbb{A}}}\uparrow & & \uparrow{\scriptstyle s} \\ \widehat{\mathbb{Z}}^{(\Sigma)\times} & \xrightarrow[\iota]{\sim} & \mathrm{Gal}(\mathbb{Q}^{(\Sigma)}/\mathbb{Q})\,, \end{array}$$

where ι is the Artin reciprocity law map given in Theorem 2.12.

Since $\mathbb{A}^{(\Sigma)\times}$ (resp. $\mathbb{Z}^{\times}_{(\Sigma)}\mathbb{R}^{\times}_{+}$) is a closed subgroup of $F_{\mathbb{A}}^{(\Sigma)\times}$ (resp. $\mathbb{A}^{(\Sigma)\times}$) and $\overline{\mathbb{A}^{(\Sigma)\times}O^{\times}_{(\Sigma)}F^{\times}_{\infty+}}$ is topologically generated by $\mathbb{A}^{(\Sigma)\times}$ and $\overline{O^{\times}_{(\Sigma)}F^{\times}_{\infty+}}$, we have

$$\overline{\mathbb{A}^{(\Sigma)\times}O^{\times}_{(\Sigma)}F^{\times}_{\infty+}}/\overline{O^{\times}_{(\Sigma)}F^{\times}_{\infty+}} \cong \mathbb{A}^{(\Sigma)\times}/\mathbb{Z}^{\times}_{(\Sigma)}\mathbb{R}^{\times}_{+} \cong \widehat{\mathbb{Z}}^{(\Sigma)\times}.$$

This combined with the above commutative diagram shows that

$$\mathcal{G}^{(\Sigma)}(G,X) = K \cdot s_{\mathbb{A}}(\widehat{\mathbb{Z}}^{(\Sigma)\times}) = s_{\mathbb{A}}(\widehat{\mathbb{Z}}^{(\Sigma)\times})K$$

for $K = \left\{ s \in G(\mathbb{A}^{(\Sigma)}) \middle| \det(s) \in \overline{O^{\times}_{(\Sigma)}F^{\times}_{\infty+}} \right\}$. Taking the finite part, we find

$$K/K_{\infty} \cong K^{(\infty)} = \left\{ s \in G(\mathbb{A}^{(\Sigma\infty)}) \middle| \det(s) \in \overline{O^{\times}_{(\Sigma)+}} \right\},$$

where $O^{\times}_{(\Sigma)+} = O_{(\Sigma)} \cap F^{\times}_{+}$. Then the strong approximation theorem (e.g., [MFG] Theorem 3.2) shows that $\mathcal{G}^{(\Sigma)}(G,X) = \overline{G(\mathbb{Z}_{(\Sigma)})_{+} \cdot s_{\mathbb{A}}(\widehat{\mathbb{Z}}^{(\Sigma)\times})}$. As seen in (4.2), for $z \in \mathfrak{Z}$, the test object $(\mathcal{L}_z, \det(\alpha)\lambda_z, \phi_{N,z} \circ \alpha^{(\infty)})$ is prime-to-Σ isogenous to $(\mathcal{L}_{\alpha^{-1}(z)}, \lambda_{\alpha^{-1}(z)}, \phi_{\alpha^{-1}(z)})$ for an element $\alpha \in G(\mathbb{Z}_{(\Sigma)})_{+} = G(\mathbb{Z}_{(\Sigma)}) \cap G(\mathbb{R})_{+}$. Thus $\alpha^{(\infty)}$ $(\alpha \in G(\mathbb{Z}_{(\Sigma)})_{+})$ sends a point in $\mathfrak{M}^{(\Sigma)}(\mathbb{C})$ represented by $z \in \mathfrak{Z}$ to that represented by $\alpha^{-1}(z)$, and hence the image of $G(\mathbb{Z}_{(\Sigma)})_{+}$ in $\overline{\mathcal{E}}(G,X)$ preserves $\mathfrak{M}^{(\Sigma)}$, in particular $G(\mathbb{Z}_{(\Sigma)})_{+} \subset \mathcal{G}^{(\Sigma)}(G,X)$ preserves the component $\mathfrak{M}^{(\Sigma)}$. We conclude from the continuity that $\tau : \overline{\mathcal{E}}^{(\Sigma)}(G,X) \hookrightarrow \mathrm{Aut}(\mathfrak{M}^{(\Sigma)}_{/\mathbb{Q}})$. Since $G(\mathbb{Q}_{\Sigma}) = B(\mathbb{Z}_{\Sigma})G(\mathbb{Z}_{\Sigma})$ for the upper triangular Borel subgroup $B \subset G$, it is easy to see that the normalizer of $G(\mathbb{Z}_{\Sigma})$ is itself. Thus any element $g \in G(\mathbb{A}^{(\infty)})$ preserving $Sh(G,X)/G(\mathbb{Z}_{\Sigma})$ has its Σ-component g_{Σ} in $G(\mathbb{Z}_{\Sigma})$, and $\overline{\mathcal{E}}^{(\Sigma)}(G,X)$ gives the stabilizer of $\mathfrak{M}^{(\Sigma)}$ in $\frac{G(\mathbb{A}^{(\Sigma\infty)})}{\overline{Z(\mathbb{Z}_{(\Sigma)})}} = \frac{G(\mathbb{Z}_{\Sigma}\times\mathbb{A}^{(\Sigma\infty)})}{\overline{Z(\mathbb{Q})\cap Z(\mathbb{Z}_{\Sigma}\times\mathbb{A}^{(\Sigma\infty)})}}$. Up to this point, the proof follows Shimura's argument (see, e.g., [AAF] Section 8).

From now on, we argue slightly differently from [Sh5] II, [Mt], and [MiS].

Step 3: Let $\underline{\mathbb{X}} = (\mathbb{X}, \overline{\lambda}, \phi)$ be the universal test object on $\mathfrak{M}^{(\Sigma)}$. If the pullback $\sigma^*\mathbb{X} = \mathbb{X} \times_{\mathfrak{M},\sigma} \mathfrak{M}$ for $\sigma \in \mathrm{Aut}(\mathfrak{M}^{(\Sigma)})$ is induced by an isogeny $\alpha : \mathbb{X} \to \sigma^*\mathbb{X}$, we find $\sigma^*\phi = \alpha \circ \phi = \phi \circ g$ for $g \in G(\mathbb{A}^{(\Sigma)})$ and hence $\tau(g) \circ \sigma = 1$, which basically finishes the proof.

Towards this goal, we introduce preparatory propositions. We start with

Proposition 4.15 *Let $\sigma \in \mathrm{Aut}(\mathfrak{K}^{(\Sigma)})$. Let $U \subset \mathfrak{M}^{(\Sigma)}$ be an open dense subscheme on which $\sigma \in \mathrm{Aut}(\mathfrak{K}^{(\Sigma)})$ induces an isomorphism $U \cong \sigma(U)$. Let*

K be an open compact subgroup of $\mathcal{E}^{(\Sigma)}(G,X)$, and write U_K for the image of U in $Sh^{(\Sigma)}/K$. Let η be the generic point of U_K and $\mathbb{X}_\eta$ be the generic fiber of the universal abelian scheme $\mathbb{X}_K$ over $\mathfrak{M}_K$. Then the two abelian varieties $\mathbb{X}_\eta$ and $\mathbb{X}_{\sigma(\eta)}$ (for the generic point $\sigma(\eta)$ of $\sigma(U_K)$) are isogenous over $\mathbb{Q}(\eta,\sigma(\eta))$ (which is the composite in $\mathfrak{K}^{(\Sigma)}$ of the two function fields of U_K and $\sigma(U_K)$).

The proof uses many facts from the theory of algebraic fundamental groups $\pi_1(X,x)$ of schemes X, which is summarized in Section 4.4. Here we use freely without explanation the terminology concerning fundamental groups, referring the reader to Section 4.4.

Proof. Open compact subgroups K of $\mathcal{E}^{(\Sigma)}(G,X)$ give a fundamental system of open neighborhoods of the identity element of $\mathrm{Aut}(\mathfrak{K}^{(\Sigma)})$ under the Krull topology. Let $\mathfrak{M}_K=\mathfrak{M}/K$ whose function field is the fixed field of K in $\mathfrak{K}^{(\Sigma)}$. Then $K=\mathrm{Gal}(\mathfrak{M}/\mathfrak{M}_K)$ for the stabilizer of $\mathfrak{M}_K$. If we choose K sufficiently small depending on $\sigma\in\mathrm{Aut}(\mathfrak{M}_{/\mathbb{Q}[\mu_N]})$, we have ${}^\sigma K=\sigma K\sigma^{-1}$. This follows from the Galois theory in Section 2.3.2. We write U_K (resp. $U_{{}^\sigma K}$) for the image of U (resp. of $\sigma(U)$) in $\mathfrak{M}_K$ (resp. in $\mathfrak{M}_{{}^\sigma K}$).

By the argument preceding this proposition, the action of the group $G(\mathbb{A}^{(\Sigma)})$ on the set $\pi_0(Sh^{(\Sigma)}_{/\overline{\mathbb{Q}}})$ of connected components factors through the determinant map of G, and the image of $G_1(\mathbb{A}^{(\Sigma)})$ in $\mathrm{Aut}(Sh^{(\Sigma)})$ is contained in the stabilizer $\mathrm{Aut}(\mathfrak{M}_{/\mathbb{Q}})$ and hence in the field automorphism group $\mathrm{Aut}(\mathfrak{K}^{(\Sigma)})$. Let $\overline{G}_1(\mathbb{A}^{(\Sigma)})$ be the image of $G_1(\mathbb{A}^{(\Sigma)})$ in $\mathrm{Aut}(\mathfrak{K}^{(\Sigma)})$ (so $\overline{G}_1(\mathbb{A}^{(\Sigma)})$ is isomorphic to the quotient of $G_1(\mathbb{A}^{(\Sigma)})$ by the center of $G_1(\mathbb{Z}_{(\Sigma)})$). We take a sufficiently small open compact subgroup S of $\mathcal{G}^{(\Sigma)}(G,X)$. We write $S_1=G_1(\mathbb{A}^{(\Sigma)})\cap S$ and $\overline{S}_1$ (resp. $\overline{S}$) for the image of S_1 (resp. S) in $\overline{G}_1(\mathbb{A}^{(\Sigma)})$ (resp. in $\overline{\mathcal{E}}(G,X)$). Shrinking S if necessary, we may assume that $\mathfrak{M}/\mathfrak{M}_S$ is étale, that the geometric Galois group $\mathrm{Gal}(\mathfrak{M}_{/\overline{\mathbb{Q}}}/\mathfrak{M}_{S/\overline{\mathbb{Q}}})\cong\mathrm{Gal}(\mathfrak{M}_{/\mathbb{Q}^{(\Sigma)}}/\mathfrak{M}_{S/\mathbb{Q}^{(\Sigma)}})=\overline{S}_1$, that $S=\prod_\ell S_\ell$ with $S_\ell=S\cap G(\mathbb{Q}_\ell)$ for primes $\ell\notin\Sigma$, and that ${}^\sigma\overline{S}_1=\sigma\overline{S}_1\sigma^{-1}\subset\overline{G}_1(\mathbb{A}^{(\Sigma)})$.

The Galois group $\mathrm{Gal}(\mathbb{Q}^{(\Sigma)}/\mathbb{Q})$ acts on the underlying topological space of $Sh^{(\Sigma)}/S$. Thus we have a finite extension k_S of $\mathbb{Q}$ inside $\mathbb{Q}^{(\Sigma)}$ such that $\mathrm{Gal}(\overline{\mathbb{Q}}/k_S)$ gives the stabilizer of $\mathfrak{M}_S$ in $Sh^{(\Sigma)}/S$. Since the action of $\mathcal{G}^{(\Sigma)}$ on the set of geometrically irreducible components $\pi_0(Sh^{(\Sigma)}_{/\overline{\mathbb{Q}}})$ factors through $\mathrm{Gal}(\mathbb{Q}^{(\Sigma)}/\mathbb{Q})$ via the determinant map composed with the Artin reciprocity map (Step 2), the field k_S is given by the fixed field of $\det(S)$ in $\mathbb{Q}^{(\Sigma)}$. The variety $\mathfrak{M}_S$ is defined over k_S. Thus $\kappa=k_S(\eta,\sigma(\eta))$ is a field of finite type over $\mathbb{Q}$; so, we can apply Theorem 4.13 to $\mathbb{X}_\eta$ and $\mathbb{X}_{\sigma(\eta)}$ defined over κ. We fix an algebraic closure $\overline{\kappa}$ of κ and take a geometric point $\overline{\eta}\in\mathfrak{M}_S(\overline{\kappa})$ over η. Let $\mathbb{X}_S$ be the universal abelian scheme over U_S. Write $\mathbb{X}_{\overline{\eta}}=\mathbb{X}_\eta\otimes_{\mathbb{Q}(\eta)}\overline{\kappa}$ for the geometric fiber of $\mathbb{X}_S$ at $\overline{\eta}$.

Note that $\pi_1(\mathbb{X}_{\overline{\eta}/\overline{\kappa}},\mathbf{0}_{\overline{\eta}})$ for the origin $\mathbf{0}_{\overline{\eta}}$ of the abelian variety $\mathbb{X}_{\overline{\eta}}$ is canonically isomorphic to the Tate module $\mathcal{T}(\mathbb{X}_{\overline{\eta}/\overline{\kappa}})$ (see [ABV] Section 18) and that $f_*\mathcal{O}_{\mathbb{X}}=\mathcal{O}_{\mathfrak{M}}$, because $\mathbb{X}\xrightarrow{f}\mathfrak{M}$ is geometrically connected (by the

construction of $\mathfrak{M}$ over $\mathbb{C}$). We have the following exact sequence by Theorem 4.43, $\mathcal{T}(\mathbb{X}_{\overline{\eta}}) \xrightarrow{\iota} \pi_1(\mathbb{X}_{S/k_S}, \mathbf{0}_{\overline{\eta}}) \to \pi_1(U_{S/k_S}, \overline{\eta}) \to 1$. This sequence is split exact, because of the zero section $\mathbf{0} : U_S \to \mathbb{X}_S$ (and covariant functoriality of fundamental groups; see Section 4.4.2). The multiplication by N: $\mathbb{X} \to \mathbb{X}$ is an irreducible étale covering, and we conclude that $\mathcal{T}^{(\Sigma)}(\mathbb{X}_x)$ injects into $\pi_1(\mathbb{X}_{S/k_S}, \mathbf{0}_{\overline{x}})$ for any geometric point $\overline{x} \in U_S$. We form the quotient $\pi_1^{\Sigma}(\mathbb{X}_{S/k_S}, \mathbf{0}_{\overline{\eta}}) = \pi_1(\mathbb{X}_{S/k_S}, \mathbf{0}_{\overline{\eta}})/i(\mathcal{T}_{\Sigma}(\mathbb{X}_{\overline{\eta}}))$, and we get a split exact sequence:

$$0 \to \mathcal{T}^{(\Sigma)}(\mathbb{X}_{\overline{\eta}}) \to \pi_1^{\Sigma}(\mathbb{X}_{S/k_S}, \mathbf{0}_{\overline{\eta}}) \to \pi_1(U_{S/k_S}, \overline{\eta}) \to 1.$$

By this exact sequence, $\pi_1(U_{S/k_S}, \overline{\eta})$ acts by conjugation on $\mathcal{T}^{(\Sigma)}(\mathbb{X}_{\overline{\eta}})$. Recall that we have chosen S sufficiently small so that $\mathfrak{M} \twoheadrightarrow \mathfrak{M}_S$ is étale. We have a canonical surjection $\pi_1(U_{S/k_S}, \overline{\eta}) \twoheadrightarrow \mathrm{Gal}(U/U_S) = \overline{S}$. Since $\mathbb{X}_{\overline{\eta}}[N]$ for all integers N outside Σ gets trivialized over U, the action of $\pi_1(U_{S/k_S}, \overline{\eta})$ on $\mathcal{T}^{(\Sigma)}(\mathbb{X}_{\overline{\eta}})$ factors through $\pi_1(U_{S/k_S}, \overline{\eta}) \twoheadrightarrow \mathrm{Gal}(U/U_S) = \overline{S}$.

We now have another split exact sequence:

$$0 \to \mathcal{T}^{(\Sigma)}(\mathbb{X}_{\sigma(\overline{\eta})}) \to \pi_1^{\Sigma}(\mathbb{X}_{{}^{\sigma}S/k_S}, \mathbf{0}_{\sigma(\overline{\eta})}) \to \pi_1(U_{{}^{\sigma}S/k_S}, \sigma(\overline{\eta})) \to 1.$$

The action of $\pi_1(U_{{}^{\sigma}S/k_S}, \sigma(\overline{\eta}))$ on $\mathcal{T}^{(\Sigma)}(\mathbb{X}_{\sigma(\overline{\eta})})$ factors through $\mathrm{Gal}(U/U_{{}^{\sigma}S}) = {}^{\sigma}\overline{S}$. We fix a path in $U_{{}^{\sigma}S}$ from $\sigma(\overline{\eta})$ to $\overline{\eta}$ (as in Remark 4.41) and lift it to a path from $\sigma(\mathbf{0}_{\overline{\eta}})$ to $\mathbf{0}_{\overline{\eta}}$ in $\mathbb{X}_{{}^{\sigma}S}$, which induces isomorphisms ([SGA] V.7):

$$\iota^{\sigma} : \pi_1^{\Sigma}(\mathbb{X}_{{}^{\sigma}S/k_S}, \mathbf{0}_{\sigma(\overline{\eta})}) \cong \pi_1^{\Sigma}(\mathbb{X}_{{}^{\sigma}S/k_S}, \mathbf{0}_{\overline{\eta}}), \quad \iota_{\sigma} : \pi_1(U_{{}^{\sigma}S}, \mathbf{0}_{\sigma(\overline{\eta})}) \cong \pi_1(U_{{}^{\sigma}S}, \overline{\eta}).$$

We want to have the following commutative diagram.

$$\begin{array}{ccccc}
\mathcal{T}^{(\Sigma)}(\mathbb{X}_{\overline{\eta}}) & \xrightarrow{\hookrightarrow} & \pi_1^{\Sigma}(\mathbb{X}_{S/k_S}, \mathbf{0}_{\overline{\eta}}) & \xrightarrow{\twoheadrightarrow} & \pi_1(U_{S/k_S}, \overline{\eta}) \\
{\scriptstyle ?}\downarrow & & \downarrow{\scriptstyle ?} & & {\scriptstyle \sigma_*}\downarrow \\
\mathcal{T}^{(\Sigma)}(\mathbb{X}_{\sigma(\overline{\eta})}) & \xrightarrow{\hookrightarrow} & \pi_1^{\Sigma}(\mathbb{X}_{{}^{\sigma}S/k_S}, \mathbf{0}_{\sigma(\overline{\eta})}) & \xrightarrow{\twoheadrightarrow} & \pi_1(U_{{}^{\sigma}S/k_S}, \sigma(\overline{\eta})) \\
{\scriptstyle \iota}\downarrow & & {\scriptstyle \iota^{\sigma}}\downarrow & & \downarrow{\scriptstyle \iota_{\sigma}} \\
\mathcal{T}^{(\Sigma)}(\mathbb{X}_{\overline{\eta}}) & \xrightarrow[\hookrightarrow]{} & \pi_1^{\Sigma}(\mathbb{X}_{{}^{\sigma}S/k_S}, \mathbf{0}_{\overline{\eta}}) & \xrightarrow[\twoheadrightarrow]{} & \pi_1(U_{{}^{\sigma}S/k_S}, \overline{\eta}).
\end{array}$$

By definition, $\mathcal{T}^{(\Sigma)}(\mathbb{X}_{\sigma(\overline{\eta})}) \xrightarrow{\iota} \mathcal{T}^{(\Sigma)}(\mathbb{X}_{\overline{\eta}})$ is an isomorphism of ${}^{\sigma}S$-modules.

We ask if we can find a linear endomorphism $\mathcal{L} \in \mathrm{End}_{\mathbb{A}^{(\Sigma)}}(\mathcal{T}^{(\Sigma)}(\mathbb{X}_{\overline{\eta}}) \otimes \mathbb{Q})$ such that $\mathcal{L}(s \cdot x) = {}^{\sigma}s \cdot \mathcal{L}(x)$ for all $s \in S$ and $x \in \mathcal{T}^{(\Sigma)}(\mathbb{X}_{\overline{\eta}})$, where ${}^{\sigma}s = \sigma s \sigma^{-1}$, and ${}^{\sigma}s$ is the image of $\iota_{\sigma}(\sigma_*(s))$ in ${}^{\sigma}S$ for any lift $s \in \pi_1(U_S, \overline{\eta})$ inducing $s \in S$. Since $\mathrm{Hom}(G_1(\mathbb{Z}_{\ell}), G_1(\mathbb{Z}_{\ell'}))$ is a singleton made of the zero-map (taking the entire $G_1(\mathbb{Z}_{\ell})$ to the identity of $G_1(\mathbb{Z}_{\ell'})$) if two primes ℓ and ℓ' are large and distinct (see (S3) in Section 4.4.3 for a proof of this fact), $s \mapsto {}^{\sigma}s$ sends $S_{1,\ell}$ into ${}^{\sigma}S_{1,\ell}$ for almost all primes ℓ, where $S_{1,\ell} = G_1(\mathbb{Q}_{\ell}) \cap S_{\ell}$. By [GAN] IV.3 (see (S1) in Section 4.4.3), if we shrink S further for exceptional finitely

many primes ℓ, we achieve that S_ℓ is ℓ-profinite, that $s \mapsto {}^\sigma s$ sends $S_{1,\ell}$ into ${}^\sigma S_{1,\ell}$, and that the isomorphism: $s \mapsto {}^\sigma s$ is induced by the conjugation of an element in $G(\mathbb{Q}_\ell)$ up to the automorphism of $\mathrm{Aut}(F_\ell/\mathbb{Q}_\ell)$: ${}^\sigma s_\ell = g_\ell s_\ell^\tau g_\ell^{-1}$ for $g_\ell \in G(\mathbb{Q}_\ell)$ and a ring automorphism τ of F_ℓ over $\mathbb{Q}_\ell$. By (S) in Section 4.4.3, the isomorphism $s \mapsto {}^\sigma s$ is induced by an element $\overline{\mathcal{L}}$ of the automorphism group $\mathrm{Aut}(G_1(\mathbb{A}^{(\Sigma\infty)}))$ fitting into the middle term of the exact sequence:

$$PG(\mathbb{A}^{(\Sigma\infty)}) \to \prod_{\ell\notin\Sigma}{}' (\mathrm{Aut}(F_\ell/\mathbb{Q}_\ell) \ltimes PG(\mathbb{Q}_\ell)) \xrightarrow{\pi_F} \mathrm{Aut}(F_{\mathbb{A}}^{(\Sigma\infty)}), \tag{4.35}$$

where $PG(\mathbb{Q}_\ell) = PGL_2(F_\ell)$. Here $\prod'$ is the restricted product with respect to S, and the semi-direct product $\mathrm{Aut}(R) \ltimes PG(R)$ (for the ring automorphism group $\mathrm{Aut}(R)$) is taken with respect to the action of $\mathrm{Aut}(R)$ on matrices entry by entry. The element $\overline{\mathcal{L}}$ in $\mathrm{Aut}(G_1(\mathbb{A}^{(\Sigma)}))$ is in turn induced by an endomorphism $\mathcal{L} \in \mathrm{End}_{\mathbb{A}^{(\Sigma)}}(\mathcal{T}^{(\Sigma)}(\mathbb{X}_{\overline{\eta}})\otimes\mathbb{Q})$. Put $g(\sigma) = \iota^{-1}\circ\mathcal{L}$. Then $g(\sigma)$ is an element of $\mathrm{Hom}_{\overline{S}_1}(\mathcal{T}^{(\Sigma)}\mathbb{X}_\eta, \mathcal{T}^{(\Sigma)}\mathbb{X}_{\sigma(\eta)})$ invertible in $\mathrm{Hom}_{\overline{S}_1}(\mathcal{T}^{(\Sigma)}\mathbb{X}_\eta, \mathcal{T}^{(\Sigma)}\mathbb{X}_{\sigma(\eta)})\otimes_{\mathbb{Z}}\mathbb{Q}$, and is S_1-linear in the sense that $g(\sigma)(sx) = {}^\sigma s \cdot g(\sigma)(x)$ for all $s \in \overline{S}_1$. As shown in Proposition 4.47, an automorphism of S_ℓ modulo center is conjugation by an element of $G(\mathbb{Q}_\ell)$ up to automorphism of $F_\ell/\mathbb{Q}_\ell$. Thus this property extends to $\overline{S}$: $g(\sigma)(sx) = \zeta({}^\sigma s \cdot g(\sigma)(x))$ for all $s \in \overline{S}$ and $\zeta \in Z(\mathbb{A}^{(\Sigma)}) \cap {}^\sigma S$. The isomorphism $\overline{S} \ni s \mapsto {}^\sigma s \in \overline{S}^\sigma$ preserves the determinant map, because the action of $G(\mathbb{A}^{(\Sigma\infty)})$ on $\pi_0(Sh^{(\Sigma)})$ factors through the determinant map. Since $s \mapsto {}^\sigma s$ preserves the determinant, we conclude $\zeta^2 = \det(\zeta) = 1$. Choosing a prime $\ell \notin \Sigma$ and shrinking S_ℓ sufficiently (so ${}^\sigma S$ also shrinks), we may assume that $\zeta_\ell = 1$ for the ℓ-component ζ_ℓ of ζ. Though $\mathcal{L}$ may depend on the choice of the path (as in Remark 4.41) from $\overline{\eta}$ to $\sigma(\overline{\eta})$, the isomorphism $g(\sigma)$ (modulo the centralizer of S_1) is independent of the choice of the path; so, we forget about the path hereafter.

Since $g(\sigma)_\ell$ is $\overline{S}$-linear, it is compatible with the action of $\mathrm{Gal}(\overline{\kappa}/\kappa)$, whose action on $\mathcal{T}_\ell\mathbb{X}_\eta$ factors through $\overline{S}$. Then by Theorem 4.13, we have

$$g(\sigma)_\ell \in \mathrm{Hom}_{\mathrm{Gal}(\overline{\kappa}/\kappa)}(\mathcal{T}_\ell\mathbb{X}_\eta, \mathcal{T}_\ell\mathbb{X}_{\sigma(\eta)}) = \mathrm{Hom}(\mathbb{X}_{\eta/\kappa}, \mathbb{X}_{\sigma(\eta)/\kappa})\otimes_{\mathbb{Z}}\mathbb{Q}_\ell,$$

and we find that $\mathbb{X}_\eta$ and $\mathbb{X}_{\sigma(\eta)}$ are isogenous over κ. By this fact, σ is induced by an isogeny $\xi : \mathbb{X}_\eta \to \mathbb{X}_{\sigma(\eta)}$, and hence, we conclude $\zeta = 1$ and that $g(\sigma)$ is $\overline{S}$-linear. □

We have a projection $\pi_F : \mathrm{Aut}(\overline{G}_1(\mathbb{A}^{(\Sigma\infty)})) \twoheadrightarrow \mathrm{Aut}_{\mathrm{alg}}(F_{\mathbb{A}}^{(\Sigma\infty)}/\mathbb{A}^{(\Sigma\infty)})$ (as in (4.35)) whose kernel is given by $PG(\mathbb{A}^{(\Sigma\infty)})$ (as shown in Section 4.4.3). By the above proof, $\sigma \in \mathrm{Aut}(\mathfrak{K}^{(\Sigma)})$ has projection to $\mathrm{Aut}_{\mathrm{alg}}(F_{\mathbb{A}}^{(\Sigma\infty)}/\mathbb{A}^{(\Sigma\infty)})$ which is written as $\sigma_F = \pi_F(\sigma)$.

Corollary 4.16 *If $\sigma \in \mathrm{Aut}(\mathfrak{K}^{(\Sigma)})$, we have $\sigma_F \in \mathrm{Aut}(F)$, where $\mathrm{Aut}(F)$ is diagonally embedded into $\prod_{\ell\notin\Sigma}\mathrm{Aut}(F_\ell) = \mathrm{Aut}(F_{\mathbb{A}}^{(\Sigma\infty)}/\mathbb{A}^{(\Sigma\infty)})$.*

Proof. By the proof of Proposition 4.15, $g(\sigma)$ (modulo $\overline{S}$) is induced by an isogeny $\xi \in \mathrm{Hom}_O(\mathbb{X}_\eta, \mathbb{X}_{\sigma(\eta)})$. Since O-linear isogenies $\mathbb{X}_\eta \to \mathbb{X}_{\sigma(\eta)}$ (keeping

the polarization) give rise to the identity coset of $\mathrm{Aut}(F) \ltimes PG(\mathbb{A}^{(\Sigma\infty)})$ in $\mathrm{Aut}(G_1(\mathbb{A}^{(\Sigma\infty)}))$ (only depending on $i : O \hookrightarrow \mathrm{End}(\mathbb{X}_{\sigma(\eta)})$ with ambiguity modulo right multiplication by $\mathrm{Aut}(F)$), we conclude $\sigma_F \in \mathrm{Aut}(F)$. □

Here is the last step.

Step 4: Replacing σ by $\sigma\sigma_F^{-1}$ for $\sigma_F \in \mathrm{Aut}(F) \subset \mathrm{Aut}(\mathfrak{K})$, we may assume that $\sigma_F = \mathrm{id}_F$ for the starting $\sigma \in \mathrm{Aut}(\mathfrak{K}^{(\Sigma)})$. Thus we may assume that $g(\sigma)S = \alpha_\eta S$ for an isogeny $\alpha_\eta : \mathbb{X}_\eta \to \mathbb{X}_{\sigma(\eta)}$. Writing the level structure on $\mathbb{X}_\eta$ as $\phi_\eta^{(\Sigma)}$, we find that $\alpha_\eta \circ \phi_\eta^{(\Sigma)} = \phi_{\sigma(\eta)}^{(\Sigma)} \circ a_S^{-1}$ for $a_S \in G(\mathbb{A}^{(\Sigma)})$. In other words, $a_S^{-1}(\sigma(\eta)) = \eta$, which implies that $a_S = \sigma$ on the Zariski open dense subset U_S of $\mathfrak{M}_S$ and, hence, they are equal on the entire $\mathfrak{M}_S$. Thus σ comes from a scheme automorphism of $\mathfrak{M}$, and σ is induced by $a = \lim_S a_S \in G(\mathbb{A}^{(\Sigma\infty)})/\overline{Z(\mathbb{Z}_{(\Sigma)})}$. Since $\overline{\mathcal{E}}^{(\Sigma)}$ is the stabilizer in $G(\mathbb{A}^{(\Sigma\infty)})/\overline{Z(\mathbb{Z}_{(\Sigma)})}$ of $\mathfrak{M}^{(\Sigma)}$ by Step 2, we conclude that $a \in \overline{\mathcal{E}}^{(\Sigma)}$. □

We now study the automorphism group of the mod p fiber of $\mathfrak{M}^{(\Sigma)}$. To ensure that the special fiber is smooth, we assume that p is unramified in $F/\mathbb{Q}$ and that Σ is a proper subset of rational primes containing the prime p. Let $\mathbb{Z}_{(p)}^{(\Sigma)}$ be the valuation ring of $\mathbb{Q}^{(\Sigma)} \subset \overline{\mathbb{Q}}$ corresponding to the embedding $\iota_p : \overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_p$. Let $Sh^{(\Sigma)}(G,X)_{/\mathbb{Z}_{(p)}^{(\Sigma)}} = \varprojlim_K Sh_K^{(\Sigma)}(G,X)$ (for K running open compact subgroups maximal at Σ) be the Kottwitz model of the Shimura variety representing the functor $\mathcal{E}_1^{(\Sigma)}$ (classifying triples $(A, \overline{\lambda}, \eta^{(\Sigma)})$ up to prime-to-Σ isogenies over $\mathbb{Z}_{(p)}^{(\Sigma)}$). Let $\mathfrak{M}^{(\Sigma)}_{/\mathbb{Z}_{(p)}^{(\Sigma)}} = \varprojlim_{p\nmid N} \mathfrak{M}(O, \Gamma(N))_{/\mathbb{Z}_{(p)}^{(\Sigma)}}$. Then the generic fiber of $\mathfrak{M}^{(\Sigma)}_{\mathbb{Q}^{(\Sigma)}}$ gives a geometrically irreducible component of the generic fiber $Sh^{(\Sigma)}(G,X)_{/\mathbb{Q}^{(\Sigma)}}$ of $Sh^{(\Sigma)}(G,X)_{/\mathbb{Z}_{(p)}^{(\Sigma)}}$. If $Y' \xrightarrow{f} Y$ is a projective morphism of Noetherian schemes with $f_*\mathcal{O}_{Y'} = \mathcal{O}_Y$, then by Zariski's connectedness theorem (e.g., [ALG] III.11.3), each special fiber $f^{-1}(y)$ is connected. Taking the smooth projective toroidal compactification $M^{(\Sigma)}_{/\mathbb{Z}_{(p)}^{(\Sigma)}}$ of $\mathfrak{M}^{(\Sigma)}_{/\mathbb{Z}_{(p)}^{(\Sigma)}}$, we find that the generic fiber $M^{(\Sigma)}_{/\mathbb{Q}^{(\Sigma)}}$ is geometrically irreducible because $M^{(\Sigma)}(\mathbb{C})$ is the toroidal compactification of $\Gamma(O,N)\backslash\mathfrak{Z}$ (which is connected). This shows $\pi_*\mathcal{O}_{M^{(\Sigma)}} = \mathbb{Q}^{(\Sigma)}$ for $\pi : M^{(\Sigma)}_{/\mathbb{Q}^{(\Sigma)}} \to Spec(\mathbb{Q}^{(\Sigma)})$. Since $\mathbb{Z}_{(p)}^{(\Sigma)}$ is normal, we find that $\pi_*\mathcal{O}_{M^{(\Sigma)}} = \mathbb{Z}_{(p)}^{(\Sigma)}$ for $\pi : M^{(\Sigma)}_{/\mathbb{Z}_{(p)}^{(\Sigma)}} \to Spec(\mathbb{Z}_{(p)}^{(\Sigma)})$ (see [ALG] III.8.8). Then by Zariski's connectedness theorem, $M^{(\Sigma)}_{/\mathbb{F}} = M^{(\Sigma)} \otimes_{\mathbb{Z}_{(p)}^{(\Sigma)}} \mathbb{F}$ remains irreducible (for the residue field $\mathbb{F} \subset \overline{\mathbb{F}}_p$ of $\mathbb{Z}_{(p)}^{(\Sigma)}$), and $M^{(\Sigma)}_{/\mathbb{F}}$ is the toroidal compactification of $\mathfrak{M}^{(\Sigma)}_{/\mathbb{F}} = \mathfrak{M}^{(\Sigma)} \otimes_{\mathbb{Z}_{(p)}^{(\Sigma)}} \mathbb{F}$. This shows that $\mathfrak{M}^{(\Sigma)}_{/\mathbb{F}}$ is irreducible. Thus $\mathfrak{M}^{(\Sigma)}_{/\mathbb{F}}$ gives a geometrically irreducible component of the special fiber $Sh^{(\Sigma)}(G,X)_{/\mathbb{F}}$ of $Sh^{(\Sigma)}(G,X)_{/\mathbb{Z}_{(p)}^{(\Sigma)}}$. In this sense, we call $\mathfrak{M}^{(\Sigma)}_{/\mathbb{Z}_{(p)}^{(\Sigma)}}$

the geometrically irreducible component of $Sh^{(\Sigma)}(G,X)_{/\mathbb{Z}^{(\Sigma)}_{(p)}}$, which contains the image of the connected component $X^+ \times 1$ of $X \times G(\mathbb{A}^{(\infty)})$. Define

$$\mathcal{D}^{(\Sigma)}(G,X) = \left\{ x \in G(\mathbb{A}^{(\Sigma\infty)}) \middle| \det(x) \in \overline{(p^{\mathbb{Z}})O^{\times}_{(p)}F^{\times}_{\infty+}} \right\},$$

whose image in $\overline{\mathcal{E}}^{(\Sigma)}(G,X)$ is given by

$$\overline{\mathcal{D}}^{(\Sigma)}(G,X) := \mathcal{D}^{(\Sigma)}(G,X)/\overline{G(\mathbb{R})_+ O^{\times}_{(\Sigma)}}.$$

A mod p version of the above proof of Theorem 4.14 gives (see [H03b] for a detailed proof):

Theorem 4.17 *Let the notation be as above. Suppose that $p \in \Sigma$ and that p is unramified in $F/\mathbb{Q}$. Write $\mathfrak{k}_p^{(\Sigma)}$ for the function field of $\mathfrak{M}^{(\Sigma)}$ over the residue field $\mathbb{F}$ of $\mathbb{Z}^{(\Sigma)}_{(p)}$. Then τ in* Theorem 4.14 *induces the following isomorphism,*

$$\mathrm{Aut}(F) \ltimes \overline{\mathcal{D}}^{(\Sigma)}(G,X) \cong \mathrm{Aut}(\mathfrak{M}^{(\Sigma)}_{/\mathbb{F}_p}) \cong \mathrm{Aut}(\mathfrak{k}_p^{(\Sigma)}).$$

In particular, the decomposition group inside $\mathrm{Aut}(\mathfrak{K}^{(\Sigma)}) = \overline{\mathcal{E}}^{(\Sigma)}(G,X)$ of the generic point of the special fiber $\mathfrak{M}^{(\Sigma)}_{/\mathbb{F}_p}$ is given by $\overline{\mathcal{D}}^{(\Sigma)}(G,X) \subset \overline{\mathcal{E}}^{(\Sigma)}(G,X)$, and we may normalize the isomorphism so that each element $\pi \in \overline{\mathcal{D}}^{(\Sigma)}(G,X)$ with $\det(\pi) = p \in F^{(\Sigma)\times}_{\mathbb{A}}$ acts on $\mathbb{F}$ by the pth power map (so that $\overline{\mathcal{D}}^{(\Sigma)}(G,X)$ acts on $\mathfrak{M}^{(\Sigma)}_{/\mathbb{F}_p}$ from the right).

4.2.3 Local Reciprocity Law

Since $Sh(G,X)(\mathbb{C}) = G(\mathbb{Q})\backslash \left(X \times G(\mathbb{A}^{(\infty)})\right)/\overline{Z(\mathbb{Q})}$, we write $[x,g] \in Sh(\mathbb{C})$ for the image of $(x,g) \in X \times G(\mathbb{A}^{(\infty)})$. A point $[x,g]$ is called a *CM point* if h_x restricted to $\mathbb{G}_m \times 1 \subset \mathbb{S}_{/\mathbb{C}} = \mathbb{G}_m \times \mathbb{G}_m$ has an image in a maximal torus $T = T_x \subset G = \mathrm{Res}_{F/\mathbb{Q}} GL(2)$ defined over $\mathbb{Q}$. Assume $[x,g]$ to be a CM point. Since $T_x(\mathbb{R})$ fixes the point $x \in X$, $(T_x/Z)(\mathbb{R})$ is in the maximal compact subgroup fixing $x \in X$; so, it is an anisotropic torus, and $T_x(\mathbb{R}) \cap G_1(\mathbb{R})$ is compact. Therefore $T_x = \mathrm{Res}_{M_x/\mathbb{Q}}\mathbb{G}_m$ for a totally imaginary quadratic extension M_x of F inside $\overline{\mathbb{Q}}$ (a *CM field* over F). In this way, the morphism $h_x : \mathbb{S}_{/\mathbb{R}} \hookrightarrow G_{/\mathbb{R}}$ defined over $\mathbb{R}$ gives rise to an embedding $\rho_x : T_x \hookrightarrow G_{/\mathbb{Q}}$ defined over $\mathbb{Q}$. Identify X^+ with $\mathfrak{Z}$ and write $z \in \mathfrak{Z}$ for the point corresponding to x via the embedding $i_\infty : \overline{\mathbb{Q}} \hookrightarrow \mathbb{C}$ (we fixed in the introduction). We choose F and M_x in $\overline{\mathbb{Q}}$. Since $\rho_x(\alpha)(z) = z$, we find

$$\rho_x(\alpha)^\iota \left(\begin{smallmatrix} z & \overline{z} \\ 1 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} \rho_x(\alpha)^{-1}(z) & \rho_x(\alpha)^{-1}(\overline{z}) \\ 1 & 1 \end{smallmatrix}\right) \left(\begin{smallmatrix} cz+d & 0 \\ 0 & c\overline{z}+d \end{smallmatrix}\right) = \left(\begin{smallmatrix} z & \overline{z} \\ 1 & 1 \end{smallmatrix}\right) \left(\begin{smallmatrix} cz+d & 0 \\ 0 & c\overline{z}+d \end{smallmatrix}\right),$$

where $\rho_x(\alpha)^\iota = \left(\begin{smallmatrix} * & * \\ c & d \end{smallmatrix}\right) \in GL_2^+(F_\infty) = G(\mathbb{R})_+$. This shows that $\varphi(cz+d) = \alpha$ in M_x for $\varphi = i_\infty|_{M_x}$, that ρ_x is the regular representation given by

$$\rho_x(\alpha)^\iota \left(\begin{smallmatrix} z & \overline{z} \\ 1 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} z & \overline{z} \\ 1 & 1 \end{smallmatrix}\right) \left(\begin{smallmatrix} \alpha & 0 \\ 0 & \overline{\alpha} \end{smallmatrix}\right),$$

and that $M_x = F[z]$. The abelian variety $\mathbb{X}_x$ over x has an analytic uniformization $\mathbb{X}_x(\mathbb{C}) = F_{\mathbb{C}}/\mathcal{L}_z$ ($F_{\mathbb{C}} = \mathbb{C}^I$) for an O-lattice $\mathcal{L}_z = \mathfrak{b}z + \mathfrak{a}^* \subset M_x$. Thus $\mathrm{End}_F^{\mathbb{Q}}(\mathbb{X}_x) \supset M_x$ and we may identify $V(\mathbb{X}_x)$ with $M_x \otimes_{\mathbb{Q}} \mathbb{A}^{(\infty)}$.

For the moment, simply write x for $[x, 1]$. Then the level structure η_x of $\mathbb{X}_x$ is given by $V(\mathbb{A}^{(\infty)}) \ni (a, b) \mapsto -a + bz \in M_x \otimes_{\mathbb{Q}} \mathbb{A}^{(\infty)}$. In particular, $\alpha \circ \eta_x = \eta_x \circ \rho_x(\alpha)$. Since $\det(\rho_x(\alpha)) = \alpha\alpha^c \in F$ for $\alpha \in M_x$ for complex conjugation c generating $\mathrm{Gal}(M_x/F)$, we find $\rho_x(\alpha) \in \mathcal{G}(G, X)$. Let $(\mathbb{X}, \overline{\boldsymbol{\lambda}}, \boldsymbol{\eta})_{/Sh(G,X)}$ be the universal triple, and $f : Sh(G, X) \to Sh(G, X)$ be an automorphism defined over $\mathbb{Q}$ fixing x. Thus f fixes the geometrically irreducible component containing $x \in Sh(G, X)$. By Theorem 4.14, we find $g \in \mathcal{G}(G, X)$ such that $f = \tau(g)$. By the universality of $\mathbb{X}$, we have a triple $(\mathcal{Y}, \overline{\boldsymbol{\lambda}}', \boldsymbol{\eta}')_{/Sh(G,X)}$ and an F-linear isogeny $a : \mathcal{Y} \to f^*\mathbb{X}$ giving an isomorphism $(\mathcal{Y}, \overline{\boldsymbol{\lambda}}', \eta') \cong (f^*\mathbb{X}, f^*\overline{\boldsymbol{\lambda}}, f^*\boldsymbol{\eta})$ in the category $\mathcal{A}_F^{\mathbb{Q}}$, where $f^*\mathbb{X} = \mathbb{X} \times_{Sh(G,X),f} Sh(G, X)$ and $f^*\overline{\boldsymbol{\lambda}}$ and $f^*\boldsymbol{\eta}$ are the pullback of $\overline{\boldsymbol{\lambda}}$ and $\boldsymbol{\eta}$. On the other hand, by the definition of the action of $\tau(g)$, we know that $(f^*\mathbb{X}, f^*\overline{\boldsymbol{\lambda}}, f^*\boldsymbol{\eta}) \cong (\mathbb{X}, \overline{\boldsymbol{\lambda}}, \boldsymbol{\eta} \circ g)$. Thus the isogeny a induces an endomorphism $a_x \in \mathrm{End}_F^{\mathbb{Q}}(\mathbb{X}_x)$, and $a_x \circ \boldsymbol{\eta}_x = \boldsymbol{\eta}_x \circ \rho_x(a_x) = \boldsymbol{\eta}_x \circ g$. This shows that $M_x = \mathrm{End}_F^{\mathbb{Q}}(\mathbb{X}_x)$ and $g = \rho_x(a_x)$. In other words, the stabilizer of $[x, 1]$ in $\mathcal{G}(G, X)$ is given by the image $T_x(\mathbb{Q})$ of $M_x^\times$.

By the de Rham comparison isomorphism combined with the Hodge decomposition (e.g., [ABV] Chapter 1), we have an M_x-linear isomorphism:

$$H^1(\mathbb{X}_x(\mathbb{C}), \mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{C} \cong H^1(\mathbb{X}_x(\mathbb{C}), \mathbb{C}) \cong H^0(\mathbb{X}_x, \Omega_{\mathbb{X}_x/\mathbb{C}}) \oplus \overline{H^0(\mathbb{X}_x, \Omega_{\mathbb{X}_x/\mathbb{C}})},$$

where $\overline{H^0(\mathbb{X}_x, \Omega_{\mathbb{X}_x/\mathbb{C}})}$ is the complex conjugation of $H^0(\mathbb{X}_x, \Omega_{\mathbb{X}_x/\mathbb{C}})$. Since $H^0(\mathbb{X}_x(\mathbb{C}), \Omega_{\mathbb{X}_x/\mathbb{C}})$ is $F \otimes_{\mathbb{Q}} \mathbb{C}$-free of rank 1 (rm4), we find $a_x + a_x^c \in F$. Since a_x preserves the polarization $\boldsymbol{\lambda}_x$ (by the expression of the polarization on λ_z for $z \in \mathfrak{Z}$), we find $a_x a_x^c \in F_+^\times$. This gives a proof of the fact that $F[a_x]$ is a CM quadratic extension, and $\mathbb{X}_x$ has complex multiplication by M_x.

Conversely, start with a test object $(A, \lambda, \eta)_{/\mathbb{C}}$ with complex multiplication by a CM field M. We find a lattice $\mathcal{L}_z = 2\pi i(\mathfrak{c}^{-1}z + O^*)$ in $F_{\mathbb{C}} = F \otimes_{\mathbb{Q}} \mathbb{C} = \mathbb{C}^I$ such that $(A, \overline{\lambda})_{/\mathbb{C}}$ is given by $(F_{\mathbb{C}}/\mathcal{L}_z, \overline{\lambda}_z)$. The action of $\alpha \in M$ on the Lie algebra $Lie(A(\mathbb{C})) = \mathbb{C}^I = F_{\mathbb{C}}$ is given by multiplication by a diagonal matrix $\mathrm{diag}[\alpha^{\varphi_1}, \ldots, \alpha^{\varphi_g}]$ for field embeddings $\varphi_j : M \hookrightarrow \mathbb{C}$ whose restriction gives all elements of I. The set $\Phi = \{\varphi_j\}$ is called the CM type of A, and $\Phi \sqcup \Phi c$ gives all embeddings of M into $\mathbb{C}$ for the generator c of $\mathrm{Gal}(M/F)$. By changing Φ by Φc if necessary (and permuting φ_j), we may assume that $i_\infty|_M = \varphi_1$, which we write simply as φ. We write $\Phi(\alpha)$ for the diagonal matrix $\mathrm{diag}[\alpha^{\varphi_1}, \ldots, \alpha^{\varphi_g}]$. We extend Φ to a representation of $M \otimes_{\mathbb{Q}} \mathbb{R} = \mathbb{C}^\Phi = \mathbb{C}^I$, where the identification $M \otimes_{\mathbb{Q}} \mathbb{R} = \mathbb{C}^\Phi$ is given by $m \otimes r \mapsto (\varphi(m)r)_{\varphi \in \Phi}$. Since $\mathcal{L}_z \otimes_{\mathbb{Z}} \mathbb{Q}$ is a vector space over M under the multiplication by $\Phi(\alpha)$, it has to be dimension 1; so, we find that $\mathcal{L}_z \subset (M \otimes 1) \subset (M \otimes_{\mathbb{Q}} \mathbb{R}) = \mathbb{C}^\Phi$. Then $M = F[z^\varphi]$.

We define $\eta_x : V(\mathbb{A}^{(\infty)}) \cong V(A) = \mathcal{L}_x \otimes_O F_{\mathbb{A}}^{(\infty)} = M_{\mathbb{A}}^{(\infty)}$ (for $x = [z,1]$) by $(a,b) \mapsto -a + bz$. Then we find $g \in G(\mathbb{A}^{(\infty)})$ so that $\eta_x \circ g = \eta$. We first assume that $g = 1$. We have a representation $\rho_x : M \hookrightarrow M_2(F)$ given by $\alpha\eta_x = \eta_x \circ \rho_x(\alpha)$; in other words, ρ_x is the regular representation $\begin{pmatrix} \alpha z^{\varphi} \\ \alpha \end{pmatrix} = \rho_x(\alpha)^{\iota} \begin{pmatrix} z^{\varphi} \\ 1 \end{pmatrix}$. By our construction, an endomorphism $\alpha : A \to A$ ($\alpha \neq 0$) induces $(A, \lambda, \eta_x) \cong (A, (\alpha\alpha^c)\lambda, \eta_x \circ \rho_x(\alpha))$, and the point $x \in Sh(G,X)(\mathbb{C})$ corresponding to (A, λ, η_x) is fixed by the action of $\tau(\rho_x(\alpha))$. Thus we have an embedding $\rho_x : T_M = \mathrm{Res}_{M_x/\mathbb{Q}}\mathbb{G}_m \hookrightarrow G(\mathbb{A}^{\infty})$ uniquely associated with the abelian scheme (A, λ, η). When $g \neq 1$, the embedding $\rho : T_M \hookrightarrow G(\mathbb{A}^{\infty})$ is given by $t \mapsto \rho_x^g(t) = g^{-1}\rho_x(t) \cdot g$ and ρ is characterized by the fact that $\tau(\rho(M^{\times}))$ fixes the point of $Sh(G,X)$ associated with (A, λ, η).

Returning to $(\mathbb{X}_x, \overline{\boldsymbol{\lambda}}_x, \boldsymbol{\eta}_x)$, we thus find that $F[a_x] = M_x$ and $M_x = F[z^{\varphi}]$ for the point $z \in \mathfrak{Z}$ corresponding to x, since the commutative subalgebra of $\mathrm{End}_F^{\mathbb{Q}}(\mathbb{X}_x)$ of dimension $2\dim(\mathbb{X}_x)$ is unique (see [ACM] Section 5 and [ABV] Section 20). When $g \neq 1$, $(A, \overline{\lambda}, \eta)$ is sitting over the CM point $[x,g] = \tau(g)([x,1])$ because $\eta = \eta_x \circ g$. In this case, we take $\rho = \rho_x^g = g^{-1}\rho_x \cdot g$ in place of ρ_x, and all the arguments work well for ρ in place of ρ_x. We have proven

Theorem 4.18 *Let $[x,g]$ be a CM point. Then there exist a CM quadratic extension M_x/F and $h_x : \mathbb{S} \hookrightarrow G$ defined over $\mathbb{R}$ which has an image in $\rho_x(T_x(\mathbb{R}))$ for a unique morphism $\rho_x : T_x = \mathrm{Res}_{M_x/\mathbb{Q}}\mathbb{G}_m \hookrightarrow G$ defined over $\mathbb{Q}$. Taking the conjugate $\rho_x^g = g^{-1}\rho_x \cdot g$, the action of $\rho_x^g(M^{\times})$ on $Sh(G,X)$ fixes the geometric point $[x,g] \in Sh(G,X)(\mathbb{C})$, and the image of $\rho_x^g(M^{\times})$ in $\frac{G(\mathbb{A}^{(\infty)})}{\overline{Z(\mathbb{Q})}}$ gives the stabilizer of the geometric point $[x,g]$. Conversely, starting a triple $(A, \lambda, \eta)_{/\mathbb{C}}$ with complex multiplication by a CM field $M \subset \mathrm{End}_F^{\mathbb{Q}}(A)$, we find a unique CM point $[x,g] \in Sh(G,X)(\mathbb{C})$ such that the fiber $(\mathbb{X}_{[x,g]}, \overline{\boldsymbol{\lambda}}_{[x,g]}, \boldsymbol{\eta}_{[x,g]})$ of the universal triple $(\mathbb{X}, \overline{\boldsymbol{\lambda}}, \boldsymbol{\eta})_{/Sh(G,X)}$ at $[x,g]$ is isomorphic to $(A, \overline{\lambda}, \eta)$. In this case, we have $M = M_x$ and the image of $T_x(\mathbb{Q})$ in $\mathrm{Aut}(Sh(G,X))$ is isomorphic to $T_x^1 = \mathrm{Ker}(T_x \xrightarrow{norm} T)$, where $T = \mathrm{Res}_{F/\mathbb{Q}}\mathbb{G}_m$.*

The last statement comes from the fact that $T_x/Z \cong T_x^1$ by $\alpha \mapsto \alpha^{1-c}$, and $\rho_x(T_x(\mathbb{Q}))/Z(\mathbb{Q}) \cong \rho_x(T_x(\mathbb{Q}))\overline{Z(\mathbb{Q})}/\overline{Z(\mathbb{Q})}$.

We write $O_M = O_x$ for the integer ring of $M = M_x$. We start with a triple $(A, \overline{\lambda}, \eta)_{/\mathbb{C}}$ of CM type (M, Φ) which is sitting on the CM point $[x,g] \in Sh(G,X)(\mathbb{C})$. We first show that we can descend the abelian variety $A_{/\mathbb{C}}$ to a number field M'' in $\overline{\mathbb{Q}}$. This fact is first proven by Shimura, Taniyama, and Weil via their theory of complex multiplication (see [ACM] 8.5). We prove this using the representability of $\mathcal{E}'_{\mathfrak{c},\Gamma(N)}$ by $\mathfrak{M}'(\mathfrak{c}, \Gamma(N))$. Suppose that $O_M \hookrightarrow \mathrm{End}_O(A)$. Then $A(\mathbb{C}) \xrightarrow[\xi]{\sim} M \otimes_{\mathbb{Q}} \mathbb{R}/(\mathfrak{a} \otimes 1)$ for an ideal $\mathfrak{a} \subset M$. The isomorphism class of A only depends on $(\mathfrak{a}, \Phi)$, and for $\sigma \in \mathrm{Aut}(\mathbb{C})$, $A^{\sigma} = A \times_{\mathbb{C},\sigma^{-1}} \mathbb{C}$ has CM type $(M, \sigma\Phi)$ because we have $\Omega^{\sigma}_{A/\mathbb{C}} = \Omega_{A^{\sigma}/\mathbb{C}}$ on which M acts by $\sigma\Phi$. Define the reflex field M' of (M, Φ) by the fixed field of $\{\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \,|\, \sigma\Phi \cong \Phi\}$. Since the composite $\mathbb{Q}^{CM}$ of all CM fields is characterized by the fact that its complex conjugation c on $\mathbb{Q}^{CM}$ is in

the center of $\mathrm{Gal}(\mathbb{Q}^{CM}/\mathbb{Q})$ and hence is unique (see [IAT] Proposition 5.11 and [ACM] 18.2), we see easily that M' is again a CM field. Thus for any $\sigma \in \mathrm{Aut}(\mathbb{C}/M')$, the isomorphism class of A^σ over $\mathbb{C}$ only depends on $\mathfrak{a}$. If $\lambda : {}^tA \cong A \otimes \mathfrak{c}$, then $\lambda^\sigma : {}^tA^\sigma \cong A^\sigma \otimes \mathfrak{c}$; so, if $(A, \lambda, \phi_N) \in \mathcal{E}'_{\mathfrak{c},\Gamma(N)}(\mathbb{C})$, its conjugate is also in $\mathcal{E}'_{\mathfrak{c},\Gamma(N)}(\mathbb{C})$. Since $\mathrm{Pic}_{A/K}$ is well-defined over K if A is defined over a subfield $K \subset \mathbb{C}$, $\mathrm{Aut}(\mathbb{C}/K)$ permutes the component group $NS(A)(K)$ of $\mathrm{Pic}_{A/K}$. Since the Néron–Severi group $NS(A)(\mathbb{C})$ is in bijection with the set of Riemann forms inside $\mathrm{Hom}_{\mathbb{Z}}(\mathfrak{a} \wedge_{\mathbb{Z}} \mathfrak{a}, \mathbb{Z})$ (see, e.g., [ABV] Section 2), $NS(A)(\mathbb{C})$ is a discrete torsion-free group with countably many elements. Since A is projective, it is defined over a field K generated over $\mathbb{Q}$ by finitely many elements (giving the coefficients of the homogeneous equations of A). Thus if $x \in NS(A)(\mathbb{C})$ is transcendental over K, its conjugates under $\mathrm{Aut}(\mathbb{C}/K)$ are uncountable, contradicting the countability of $NS(A)(\mathbb{C})$. Thus a polarization λ is always defined over a finite extension of K. Therefore the number of isomorphism classes of (A, λ, ϕ_N) over $\mathbb{C}$ is countably many, because that of A is bounded by the class number of M times the number of CM types. This shows the number of conjugates under $\mathrm{Aut}(\mathbb{C})$ of the point $x \in \mathfrak{M}'(\mathfrak{c}, \Gamma(N))(\mathbb{C})$ carrying (A, λ, ϕ_N) is countably many; so, x has coordinates in a finite extension M''/M'. Choosing N sufficiently large so that $\mathcal{E}'_{\mathfrak{c},\Gamma(N)}$ is represented by $\mathfrak{M}'(\mathfrak{c}, \Gamma(N))$, we find that (A, λ, ϕ_N) is defined over M''; in particular, (A, λ) is defined over M''.

We let $G(\mathbb{Q})$ act on the column vector space $V = F^2$ through the matrix multiplication. The action of T_x via ρ_x^g makes $V_{\mathbb{A}}$ a free module over the adele ring $M_{\mathbb{A}}$ of rank 1. Then the subspace $V_x = V \otimes_{\mathbb{Q}} \mathbb{C}$ on which h_x acts by its restriction $\mu_x = h_x|_{\mathbb{G}_m \times 1}$ is preserved by multiplication by M_x, yielding an isomorphism class Φ of representations of M. In this way, we rediscover the CM type Φ, because V_x gives rise to the Lie algebra of $A(\mathbb{C})$ as we put the complex structure on $V_{\mathbb{R}}$ by h_x. Since the isomorphism class Φ is determined by its diagonal entries $\varphi_i : M_x \hookrightarrow \mathbb{C}$, we may identify Φ with a formal sum $\sum_i \varphi_i$ or the set $\{\varphi_1, \ldots, \varphi_g\}$ as before. Thus the field M' is the minimal field of definition of $\mu_x : \mathbb{G}_m \to G$ (the half of h_x). We define a morphism $r_x : T'_x = \mathrm{Res}_{M'/\mathbb{Q}}\mathbb{G}_m \to T_x$ by

$$r_x : T'_x = \mathrm{Res}_{M'/\mathbb{Q}}\mathbb{G}_m \xrightarrow{\mu_x} \mathrm{Res}_{M'/\mathbb{Q}}T_x \xrightarrow{Norm} T_x.$$

Since a morphism of the torus $\mathrm{Res}_{M'/\mathbb{Q}}\mathbb{G}_m$ into another torus of the form $\mathrm{Res}_{K/\mathbb{Q}}\mathbb{G}_m$ for a number field K can always be realized by $x \mapsto \det \circ \Phi'(x)$ for a representation Φ' of the algebra M' on a vector space over K, we can write $r_x = \det \circ \Phi'$. Then Φ' is a CM type of M', whose explicit form is given in [ACM] Section 8. The pair (M', Φ') is called the reflex CM type of (M, Φ).

For each $s \in T'_x(\mathbb{A}^{(\infty)}) = {M'}^{\times}_{\mathbb{A}^{(\infty)}}$, we have the Artin reciprocity image $[s, M'] \in \mathrm{Gal}(M'_{ab}/M')$, where M'_{ab} is the maximal abelian extension of M'. Since $T_x(\mathbb{R})$ is the stabilizer of x, we have $[x, \gamma g] = [\gamma^{-1}(x), g] = [x, g]$ for $\gamma \in T_x(\mathbb{Q})$, and the map: $[x, g] \mapsto [x, \rho_x(r_x(s))g] = [x, g \cdot \rho_x^g(r_x(s))]$ only depends on the Artin symbol $[s, M'] \in \mathrm{Gal}({M'}^{ab}/M')$ by class field theory.

We now recall the main theorem of complex multiplication in [ACM] 18.6.

Theorem 4.19 *Let $(A,\lambda,\eta)_{/\overline{\mathbb{Q}}}$ of CM type (M,Φ) be the triple corresponding to a CM point $[x,g]\in Sh(G,X)(\mathbb{C})$. Let $\xi : M\otimes_{\mathbb{Q}}\mathbb{R}/(\mathfrak{a}\otimes 1)\cong A(\mathbb{C})$ be a complex uniformization. Let $\sigma\in \mathrm{Aut}(\mathbb{C}/M')$ for the reflex field M' of (M,Φ). Take $s\in M'^{\times}_{\mathbb{A}}$ (with $s_\infty = 1$) corresponding to σ by the Artin reciprocity law; so, $[s,M']=\sigma$ on $M'^{ab}_{/M'}$. Then $(A^\sigma,\overline{\lambda}^\sigma,\eta^\sigma)\cong(A,\overline{\lambda},\eta\circ\rho^g_x(r_x(s^{-1})))$ in $\mathcal{A}^{\mathbb{Q}}_F$.*

Formulation of the theorem is slightly different from the one given in [ACM] 18.6, which asserts $(A^\sigma,\overline{\lambda}^\sigma,\eta^\sigma)\cong(A,\overline{\lambda},r_x(s^{-1})\eta)$. This is equivalent to our formulation by our definition of $\rho^g_x : T_x(M^{\times}_{\mathbb{A}^{(\infty)}})\hookrightarrow G(\mathbb{A}^{(\infty)})$. There is a motivic version of this theorem; see [Mi] I.5.

Shimura's reciprocity law for a CM point (see [ACM] 18.6, 18.8, and [Mi] II.5.1) follows from Theorem 4.19, since $Sh(G,X)$ represents the functor $\mathcal{E}^{(\emptyset)}_1$.

Corollary 4.20 (Reciprocity at a CM point) *Let $[x,g]$ be a CM point of CM type (M,Φ) in $Sh(G,X)(\mathbb{C})$. Then, we have $[x,g]\in Sh(G,X)(M'^{ab})$ for the maximal abelian extension $M'^{ab}\subset\mathbb{C}$ of the reflex field M' of (M,Φ), and for any $s\in T'_x(\mathbb{A}^{(\infty)})$, we have*

$$[s^{-1},M']([x,g]) = [x,g](\rho^g_x(r_x(s))) = [x,g\rho^g_x(r_x(s))] = [x,\rho_x(r_x(s))g].$$

Proof. We look at the CM test object $(A,\overline{\lambda},\eta\circ g)$ as in Theorem 4.19 sitting over the CM geometric point $[x,g]\in Sh(G,X)(\mathbb{C})$. Then for $\sigma\in\mathrm{Aut}(\mathbb{C}/M')$ inducing $[s,M']$ on M'^{ab}, we have by Theorem 4.19,

$$(A^\sigma,\overline{\lambda}^\sigma,(\eta\circ g)^\sigma)\cong(A,\overline{\lambda},r_x(s^{-1})\eta\circ g) = (A,\overline{\lambda},\eta\circ g\cdot\rho^g_x(r_x(s^{-1}))).$$

The test object at the right-hand side is sitting over $[x,g\rho^g_x(r_x(s^{-1}))]$, and the test object at the left-hand side is sitting over $\sigma([x,g])$. Thus the two geometric points $\sigma([x,g])$ and $[x,g\rho^g_x(r_x(s^{-1}))] = [x,\rho_x(r_x(s^{-1}))g]$ are equal in $Sh(G,X)(\mathbb{C})$: $\sigma([x,g]) = [x,g\rho^g_x(r_x(s^{-1}))] = [x,\rho_x(r_x(s^{-1}))g]$. In particular, the point $[x,g]$ is fixed by $\mathrm{Aut}(\mathbb{C}/M'^{ab})$ by the above expression, and we have $[x,g]\in Sh(G,X)(M'^{ab})$. Then we may replace σ in the above displayed formula by $[s,M']$, which gives the last identity in the corollary. □

4.2.4 Hilbert Modular Igusa Towers

Let $\mathcal{W}\subset\overline{\mathbb{Q}}$ be the strict Henselization of $\mathbb{Z}_{(p)}$ associated with the embedding $i_p : \overline{\mathbb{Q}}\hookrightarrow\overline{\mathbb{Q}}_p$. We look into the Igusa tower $T_\alpha = T_{1,\alpha/\overline{\mathbb{F}}_p}$ in (4.11) over the toroidal compactification M of the Hilbert–Blumenthal moduli $\mathfrak{M}$. Write $T^\circ_\alpha = T_\alpha\times_M\mathfrak{M}$. We prove the following theorem.

Theorem 4.21 (Ribet) *Let p be a prime outside N. Then the Igusa variety T°_α over $\mathfrak{M}(\mathfrak{c},\Gamma)_{/\overline{\mathbb{F}}_p}$ (and hence T_α) is irreducible for all α, where Γ is any one of $\Gamma(N)$, $\Gamma^1_1(\mathfrak{N})$, and $\Gamma_1(\mathfrak{N})$.*

This result was first shown by Ribet [Ri] (and [DR]) using the description (due to Deligne [D]) of p-ordinary AVRMs and density of ordinary points. We give a proof assuming that p is *unramified* in $F/\mathbb{Q}$. Our proof is a generalization of the proof given for the elliptic Igusa tower (Theorem 3.3), and uses the description of the automorphism group of the Hilbert modular automorphic function field (Theorem 4.14) and the existence of a smooth projective toroidal compactification of Hilbert modular varieties ([Ra] and [C]). Another proof of irreducibility has been proposed in [C] along the lines of the proof in the Siegel modular case of Chai–Faltings ([DAV] V.7). The point of Chai's argument is that the monodromy group at the cusp of the Igusa tower is isomorphic to the p-adic closure of the totally positive unit group $O_+^\times$, and therefore, if the closure is full in the norm-1-subgroup $\mathrm{Ker}(N : O_p^\times \to \mathbb{Z}_p^\times)$, the irreducibility follows. The Leopoldt conjecture predicts that the closure is of finite index in the norm-1-subgroup, but it is easy to create an example (even for a quadratic F) such that the index is nontrivial; so, Chai's proof works only in limited cases (see [DT] Section 12 for more details of this point). This type of phenomenon always occurs when we deal with G of the form $\mathrm{Res}_{F/\mathbb{Q}}G_0$ with $F \neq \mathbb{Q}$ for unitary or symplectic groups G_0 over F and, hence, the computation of the cuspidal monodromy does not yield irreducibility of the Igusa tower in general.

Proof. Let $N > 0$ be an integer prime to p, and write $S_N^\circ = \mathfrak{M}(\mathfrak{c}, \Gamma(N))[\frac{1}{E}]_{/\mathcal{W}}$ for a lift E of the Hasse invariant; so, $E \in H^0(M_N^*, \underline{\omega}^{a(p-1)I})$ for the minimal compactification M_N^* of $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ with $E \equiv H^a \mod p$ for the Hasse invariant H in Section 4.1.6. Recall that S_N° is a geometrically irreducible quasi-projective scheme over $\mathcal{W}$ (i.e., $S_N^\circ \otimes_{\mathcal{W}} \mathcal{K}$ and $S_N^\circ \otimes_{\mathcal{W}} \overline{\mathbb{F}}_p$ are both geometrically irreducible). We write $S_N^* = M_N^*[\frac{1}{E}]$.

Let $T_{N,\alpha}^\circ$ be the Igusa tower over $S_{N/\mathbb{F}}^\circ$ ($\mathbb{F} = \overline{\mathbb{F}}_p = W_1 = \mathcal{W}/p\mathcal{W}$) classifying quadruples $(A, \lambda, \phi_N, i_p)$ for $i_p : \mu_{p^\alpha} \otimes \mathfrak{d}^{-1} \hookrightarrow A$. Taking a toroidal compactification M_N of $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ and its ordinary locus $S_N = M_N[\frac{1}{E}]$, we can define the Igusa tower $T_{N,\alpha}$ over S_N in exactly the same way as $T_{N,\alpha}^\circ$ in (4.11). Since $T_{N,\alpha}$ is étale over S_N, the local structure around the étale neighborhood of a cuspidal divisor of $T_{N,\alpha}$ is the same as that of S_N; so, we can blow down $T_{N,\alpha}$ to get the minimal compactification $T_{N,\alpha}^*$ of $T_{N,\alpha}$ over S_N^*. In other words, $T_{N,\alpha}^*$ is the normalization of S_N^* in $T_{N,\alpha}$. We used in (4.11) the étale dual $\widehat{\mathcal{G}[p^\alpha]^\circ}$ to construct $T_{N,\alpha}$ over the toroidal compactification S_N. We could have used the maximal étale quotient $\mathbb{X}[p^\alpha]^{et}$ of the universal abelian scheme $\mathbb{X}$ over $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ to get $T_{N,\alpha}^\circ \subset T_{N,\alpha}$ over S_N° (removing the cuspidal divisors), since $\widehat{\mathcal{G}[p^\alpha]^\circ} \cong \mathbb{X}[p^\alpha]^{et}$ over S_N° by the polarization. By Koecher's principle, we have $H^0(T_{N,\alpha}^\circ, \underline{\omega}^{jI}) = H^0(T_{N,\alpha}^*, \underline{\omega}^{jI})$. Thus we do not really need $T_{N,\alpha}^*$ in the sequel, although we use the symbol $T_{N,\alpha}^*$ in order to have well-defined cusps on the variety. We write $T_\alpha = \varprojlim_N T_{N,\alpha}$ and $T_\alpha^? = \varprojlim_N T_{N,\alpha}^?$ for $? = *$ and $\circ$, where N runs over all integers prime to p. Similarly we define $S = \varprojlim_N S_N$ and $S^? = \varprojlim_N S_N^?$.

We have a tower of irreducible subvarieties $\{C_\alpha \subset T^\circ_\alpha\}_\alpha$ with closure C^*_α in T^*_α containing the infinity cusp. The variety C_α is an irreducible Galois covering over S°. Thus we only need to prove $T^\circ_\infty = C_\infty$. By the action on the level $\Gamma_1(p^\infty)$-structure $i_p \mapsto i_p \circ a$, $O_p^\times$ acts on T°_∞/S°, and $T^\circ_\infty/O_p^\times = S^\circ$. Thus we need to show that $\mathrm{Gal}(C_\infty/C_0) = O_p^\times$. In any case, we can identify $\mathrm{Gal}(C_\infty/C_0)$ with a subgroup H of $T(\mathbb{Z}_p) = O_p^\times$.

Over $\mathbb{Q}$, we can form a smooth geometrically irreducible toroidal compactification of $\mathfrak{M}(\mathfrak{c}, \Gamma(N))_{/\mathbb{Q}}$ for any level including N divisible by p; so, we have q-expansion at ∞ of sections of $\underline{\omega}^{jI}$ (which determines the sections over $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$), and the minimal compactification $M^*(\mathfrak{c}, \Gamma(N))_{/\mathbb{Q}}$ is given by $Proj(\mathfrak{G}_{\Gamma(N)})$, where $\mathfrak{G}_{\Gamma(N)} = \bigoplus_{j\geq 0} G_{jI}(\mathfrak{c}, \Gamma(N); \mathbb{Q}^{cyc})$. Then any element f in the function field $\mathfrak{K} = \mathbb{Q}^{cyc}(\mathfrak{M}_\mathfrak{c})$ (for the component $\mathfrak{M}_\mathfrak{c}$ as in (4.28)) can be written as $\frac{n}{d}$ with $n, d \in G_{jI}(\mathfrak{c}, \Gamma(N); \mathbb{Q}^{cyc})$ for $j \geq 2$. The minimal compactification is projective and can be embedded into a projective space using the theta series, which has integral q-expansion coefficients (cf. [CSM] and [AAF] Section 6). Thus for a given $f \in \mathfrak{K}$, we can take n and d such that $p^r n$ and $p^r d$ (for some $r > 0$) have $\mathcal{W}$-integral q-expansion coefficients. Indeed $G_{jI}(\mathfrak{c}, \Gamma(N); \mathbb{Q}^{cyc})$ has a $\mathcal{W}$-integral base for $j \geq 2$, which can be proven, for example, using the duality over $\mathbb{Q}$ of the Hecke algebra and the space of modular forms combined with the existence of the Eichler–Shimura isomorphism (e.g., [IAT] Theorem 3.52 and [H88] Sections 5 and 6). Thus, writing the q-expansion of n as $\sum_\xi a_\xi(n) q^\xi$, the valuation

$$v(n) = \mathrm{Inf}_\xi \, ord_p(a_\xi(n)) \tag{4.36}$$

is a well-defined element in $\mathbb{Q} \sqcup \{\infty\}$. Similarly $v(d) \in \mathbb{Q}$ is also well-defined. Then we define the valuation $v : \mathfrak{K} \to \mathbb{Q} \sqcup \{\infty\}$ by $v(f) = v(n) - v(d)$.

By Theorem 4.14, we have a split exact sequence,

$$1 \to \overline{\mathcal{E}}(G, X) \xrightarrow{\tau} \mathrm{Aut}(\mathfrak{K}) \to \mathrm{Aut}(F) \to 1,$$

where τ is an isomorphism of topological groups onto its image. Since the restriction of $\tau(x)$ to $\mathbb{Q}^{cyc}$ is given by the action of

$$\det(x) \in \mathbb{A}^\times/\mathbb{Q}^\times\mathbb{R}^\times_+ = \mathbb{A}^\times \overline{F^\times F^\times_{\infty+}}/\overline{F^\times F^\times_{\infty+}}$$

in Theorem 2.12, we find another more geometric exact sequence

$$1 \to G_1(\mathbb{A}^{(\infty)})/\{\pm 1\} \xrightarrow{\tau} \mathrm{Aut}(\mathfrak{M}_{\mathfrak{c}/\mathbb{Q}^{cyc}}) \to \mathrm{Aut}(F) \to 1.$$

We write $\mathcal{D}_p$ for the decomposition group in $\tau(\overline{\mathcal{E}}(G, X)) \subset \mathrm{Aut}(\mathfrak{K})$ for this valuation v. As seen in the proof of Theorem 3.3, $\mathcal{D}_p$ is a closed subgroup of $\tau(\overline{\mathcal{E}}(G, X))$, and we need to prove that $\mathcal{D}_p/I_p$ contains $O_p^\times = T(\mathbb{Z}_p)$ for the inertia subgroup I_p.

Since the quadruples $(A, \lambda, \phi_N, i_p : \mu_{p^\alpha} \otimes \mathfrak{d}^{-1} \hookrightarrow A)$ are classified over $\mathcal{W}$ by $Sh(G, X)/(S(N) \cap S^1_1(p^\alpha))$, it appears that we only need to look at the decomposition group $\mathcal{D}'_p$ of v in $\mathrm{Aut}(\mathfrak{M}_{U/\mathbb{Q}^{cyc}})$ ($\mathfrak{M}_U = \mathfrak{M}_\mathfrak{c}/U_{p/\mathbb{Q}^{cyc}}$) for

$$U_p = \bigcap_{p\nmid N,\alpha} (S(N) \cap S_1^1(p^\alpha)) = \left\{ \left(\begin{smallmatrix} 1 & u \\ 0 & 1 \end{smallmatrix}\right) \middle| u \in O_p \right\}.$$

However, we need to study $\mathcal{D}'_p$ as well as the full decomposition group $\mathcal{D}_p$.

The quotient $G_1(\mathbb{A}^{(p\infty)})/\{\pm 1\}$ acts continuously on $S^\circ_{/\mathbb{F}}$ and $S^*_{/\mathbb{F}}$, and as we remarked, S° and S^* are geometrically irreducible. Since $\mathfrak{M}_U$ and T°_∞ both classify quadruples $(A, \lambda, \eta^{(p)} : V(\mathbb{A}^{(p\infty)}) \cong V^{(p)}(A), i_p : \mu_{p^\infty} \otimes \mathfrak{d}^{-1} \hookrightarrow A[p^\infty])$ over $\mathbb{Q}^{cyc}$ and over $\overline{\mathbb{F}}_p$ and since C^*_∞ contains the infinity cusp, the function field of C_∞ is the residue field of the valuation v restricted to the function field of $\mathfrak{M}_U$ over $\mathbb{Q}^{cyc}$. The action of $a \in T(\mathbb{Z}_p)$ on the function field of C_∞ and the function field of $\mathfrak{M}_U$ is induced by the action on the level p-structure: $i_p \mapsto i_p \circ a$. Since the action of $g \in G(\mathbb{A}^{(p\infty)})$ is given by the action of the prime-to-p level structure, the action of $g \in G(\mathbb{A}^{(p\infty)})$ on S° extends to T°_∞ in the same way: $\eta^{(p)} \mapsto \eta^{(p)} \circ g$. Since the action of $T(\mathbb{Z}_p) = O_p^\times$ for the diagonal torus T of G_1 commutes with the action of $G_1(\mathbb{A}^{(p\infty)})$ and coincides with the action of $O_p^\times$ on T°_∞/S°, we have a commutative diagram with exact rows,

$$\begin{array}{ccccc} \mathrm{Gal}(C_\infty/S^\circ) & \hookrightarrow & \mathrm{Aut}(C_\infty; S^\circ) & \xrightarrow{\mathrm{Res}} & \mathrm{Aut}(S^\circ) \\ \wr\downarrow & & \uparrow\cup & & \uparrow\cup \\ H & \hookrightarrow & (G_1(\mathbb{A}^{(p\infty)})/\{\pm 1\}) \times H & \twoheadrightarrow & G_1(\mathbb{A}^{(p\infty)})/\{\pm 1\}, \end{array}$$

where $\mathrm{Aut}(C_\infty; S^\circ)$ is the group of automorphisms of C_∞ that induce automorphisms of S°. This commutative diagram shows that the action of $g \in G(\mathbb{A}^{(p\infty)})$ preserves the irreducible component $C_\infty \subset T^\circ_\infty$.

The stability of C_∞ under $G(\mathbb{A}^{(p\infty)})$ can also be proven as follows. By construction, the image $C_{N,\alpha}$ of C_α in T°_N for an integer N prime to p is a subscheme of $\mathbb{X}[Np^\alpha]^{et}$ (for the universal abelian scheme $\mathbb{X}_{/S_1}$), and the action of g with $g_p = 1$ only concerns the prime-to-p part that is canonically a factor of $\mathbb{X}[Np^\alpha]^{et}$; so, the irreducible component containing ∞ is kept under $\tau(g)$. Therefore $g \in G(\mathbb{A}^{(p\infty)})$ sends the cusp ∞ to another cusp s on the same irreducible component C_∞. This shows that $\tau(g)$ for $g \in G_1(\mathbb{A}^{(p\infty)})/\{\pm 1\} \subset \overline{\mathcal{E}}(G, X)$ preserves v.

There is one more way to show the stability of C_∞ under $G(\mathbb{A}^{(p\infty)})$. Looking at the Tate abelian variety $\mathrm{Tate}_{O,\mathfrak{c}^{-1}}(q) = \widehat{\mathbb{G}}_m \otimes \mathfrak{d}^{-1}/q^O$ with level structure $\phi_{can} : (\mu_N \otimes \mathfrak{d}^{-1}) \times (O/NO) \cong \mathrm{Tate}_{O,\mathfrak{c}^{-1}}(q)[N]$ sending $(\zeta \otimes a, m)$ to $(\zeta \otimes a)q^{m/N}$, we easily see that $\tau \left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right)^{-1}$ with $\left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right) \in SL_2(F)$ sends q-expansion $\sum_\xi a_\xi q^\xi$ of a modular form to $\sum_\xi a_\xi \exp(2\pi i \mathrm{Tr}(\frac{\xi b}{d}))q^{\xi a/d}$; so, the image of $B(\mathbb{Q}) \cap G(\mathbb{Q})^+$ (for the upper triangular Borel subgroup $B(\mathbb{Q}) \subset G(\mathbb{Q})$) in $\overline{\mathcal{E}}(G, X)$ preserves v. For the unipotent radical U of B, we find $\tau(U(\mathbb{A}^{(\infty)})) \subset \mathcal{D}_p$, because $\mathcal{D}_p$ is closed and $U(\mathbb{Q})$ is dense in $U(\mathbb{A}^{(\infty)})$. Since D_ℓ^{cusp} in Theorem 4.9 for the variety $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ is p-adically dense in V^{cusp}, we find that $V^{cusp}_{1,\infty} = D_\ell^{cusp} \otimes_W \mathbb{F}$. We can extend (using q-expansion of elements in V^{cusp}) the valuation v to D_ℓ^{cusp} and $f \in \mathfrak{K}^{U(\mathbb{Z}_p)}$ can be written as $\frac{g}{h}$ for $g, h \in D_\ell^{cusp}$ for a sufficiently large integer N prime to p. Since g and h

are defined over $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ for N prime to p, their valuations are unchanged under the action of $G_1(\mathbb{A}^{(p\infty)})$; so, the valuation v on $\mathfrak{K}^{U(\mathbb{Z}_p)}$ is preserved by $G_1(\mathbb{A}^{(p\infty)})$. As we have already seen, $\tau(U(\mathbb{Z}_p))$ preserves v on $\mathfrak{K}$, $G_1(\mathbb{A}^{(p\infty)})$ preserves v on $\mathfrak{K}$, and hence it preserves C_∞.

By the stability of v under $\tau(G_1(\mathbb{A}^{(p\infty)}))$, the closure of $\tau(G_1(\mathbb{A}^{(\infty)})B(\mathbb{Q}))$ in $\mathrm{Aut}(\mathfrak{K})$ is in $\mathcal{D}_p$. Note that the projection of $G_1(\mathbb{A}^{(\infty)})B(\mathbb{Q})$ into $G_1(\mathbb{Q}_p)$ contains $B(\mathbb{Q}) \subset G_1(\mathbb{Q}_p)$, which is dense in $B(\mathbb{Q}_p)$ (p-adic density of F in F_p). Thus $\mathcal{D}_p$ contains the image under τ of

$$T(\mathbb{Z}_p) = \left\{ \left(\begin{smallmatrix} a & 0 \\ 0 & a^{-1} \end{smallmatrix}\right) \in SL_2(O_p) \right\} \cong O_p^\times .$$

Since $T(\mathbb{Z}_p)$ normalizes U_p, the decomposition group $\mathcal{D}'_p$ of v in $\mathrm{Aut}(\mathfrak{M}_{U/\mathbb{Q}^{cyc}})$ contains the image of $T(\mathbb{Z}_p)$ isomorphic to $O_p^\times$.

As we already remarked, the action of $a \in T(\mathbb{Z}_p)$ on $\mathfrak{M}_U$ is induced by the action on test objects $(A, \lambda, \eta^{(p)}, i_p) \mapsto (A, \lambda, \eta^{(p)}, i_p \circ a)$. This shows that $\mathrm{Gal}(C_\infty/C_0) = O_p^\times$, and hence T_α is irreducible over $\overline{\mathbb{F}}_p$. □

In the above proof, we have shown $T(\mathbb{Q}_p) \subset \mathcal{D}_p$ which acts by automorphisms on $\mathfrak{M}_{\mathfrak{c}/\mathbb{Q}}$, but a nonunit $x \in T(\mathbb{Q}_p) \cap O_p$ induces only an endomorphism of the characteristic p variety T°_∞ (not an automorphism; e.g., $p \in O_p$ gives the relative Frobenius pth power map on T°_∞).

Corollary 4.22 *Let $\mathcal{D}_p \subset \tau(\overline{\mathcal{E}}(G,X))$ be the decomposition group of the valuation v defined in* (4.36). *If p is unramified in $F/\mathbb{Q}$, the morphism τ in* Theorem 4.14 *induces an isomorphism onto $\mathcal{D}_p$ from the image in $\overline{\mathcal{E}}(G,X)$ of*

$$\left\{ g \in \mathcal{G}(G,X) | g_p \in B(\mathbb{Q}_p),\ \det(g) \in \overline{\mathbb{Q}_p^\times F^\times F_{\infty+}^\times} \right\}.$$

Proof. We have already shown that $\tau(B(\mathbb{Q})) \subset \mathcal{D}_p$ and $\tau(G_1(\mathbb{A}^{(p\infty)})) \subset \mathcal{D}_p$. Since $\tau\left(\begin{smallmatrix} t & 0 \\ 0 & 1 \end{smallmatrix}\right)$ with $t \in \mathbb{A}^\times$ acts through the Artin reciprocity on the q-expansion coefficients at the infinity cusp, we find that $\tau\left(\begin{smallmatrix} t & 0 \\ 0 & 1 \end{smallmatrix}\right)$ with $t \in \mathbb{Q}_p^\times$ is in $\mathcal{D}_p$. These elements topologically generate the subgroup indicated in the corollary. Conversely, if $\tau(g) \in \mathcal{D}_p$ for $g \in \mathcal{G}(G,X)$, since $\tau(g)$ acts by $\det(g)$ on $\mathbb{Q}^{cyc}$ via the Artin reciprocity, we find that $\det(g) \in \overline{\mathbb{Q}_p^\times F^\times F_{\infty+}^\times}$. Thus modifying g by an element in $B(\mathbb{Q}_p)$, we may assume that $g_p \in GL_2(O_p)$. We consider the universal abelian scheme $\mathbb{X}_{/S^\circ}$. We then have the projection map $\pi : \mathbb{X}[p^\infty] \to \mathbb{X}[p^\infty]^{et}$. Since the irreducible component C_α is the generator of $\mathbb{X}[p^\alpha]^{et}$, looking at the infinity cusp taking the geometric fiber of $\mathrm{Tate}_{O,\mathfrak{c}^{-1}}(q)[p^\infty]$ over $\overline{\mathbb{Q}}$, g has to preserve $\mu_{p^\infty} \otimes \mathfrak{d}^{-1}$ in $\mathrm{Tate}_{O,\mathfrak{c}^{-1}}(q)[p^\infty]$; so, $g_p \in B(\mathbb{Q}_p)$. This shows the desired identity. □

We give a slightly different proof of the fact corresponding to Corollary 4.22 for Siegel modular functions in Theorem 6.28. We later further relate Corollary 4.22 to Theorem 6.28 by embedding the general Shimura variety (including Hilbert modular varieties) into the Siegel modular variety (see Theorem 8.15 and Theorem 8.16).

We now state as a corollary to Theorem 4.21 a q-expansion principle. We keep using the notation introduced in the theorem. Let $(\mathfrak{a}, \mathfrak{b})$ be a pair of fractional ideals prime to $p\mathfrak{N}$ with $\mathfrak{a}\mathfrak{b}^{-1} = \mathfrak{c}$. We have the semi-AVRM $\mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q)$ with canonical Γ-level structure i_{can} and canonical level p^∞-structure $i_{can,p}$ as in Section 4.1.5. Since $M(\mathfrak{c}, \Gamma(N))/M(\mathfrak{c}, \Gamma(1))$ for N prime to p is a finite Galois covering by the construction of the toroidal compactification, for any finite subgroup $\overline{\Gamma} \subset \mathrm{Gal}(M(\mathfrak{c}, \Gamma(N))/M(\mathfrak{c}, \Gamma(1)))$, we can make a quotient $M(\mathfrak{c}, \Gamma) = M(\mathfrak{c}, \Gamma(N))/\overline{\Gamma}$, where Γ is a symbol indicating a subgroup between $\Gamma(N)$ and $\Gamma(1)$ giving rise to $\overline{\Gamma}$. Accordingly, we have the minimal compactification $M^*(\mathfrak{c}, \Gamma) = Proj(\mathfrak{G}_\Gamma)$ for $\mathfrak{G}_\Gamma = \bigoplus_{J\geq 0} H^0(M(\mathfrak{c}, \Gamma), \underline{\omega}^{JI})$. By the same quotient process, we also have the Igusa tower $T_\alpha/M(\mathfrak{c}, \Gamma)$ and the result of Theorem 4.21 is valid in this slightly more general setting (by our construction).

For a cusp $(\mathfrak{a}, \mathfrak{b})$ of $M^*(\mathfrak{c}, \Gamma)$, we write

$$f(\mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q), \lambda_{can}, i_{can}, \omega_{can}) = a_{\mathfrak{a},\mathfrak{b}}(0, f) + \sum_{\xi\in(\mathfrak{a}\mathfrak{b})_+} a_{\mathfrak{a},\mathfrak{b}}(\xi, f)q^\xi.$$

Corollary 4.23 (q-Expansion Principle) *Let the notation be as in Theorem 4.21. Suppose that p is unramified in $F/\mathbb{Q}$, and let R be a W_m-algebra ($W_m = \mathcal{W}/p^m\mathcal{W}$). Let T_α be the Igusa tower over $M(\mathfrak{c}, \Gamma)_{/R}$. Let R' be a W_m-subalgebra of R over which T_α is well-defined. Let $f \in H^0(T_{\alpha/R}, \underline{\omega}^k_{/R}) = H^0(T^\circ_{\alpha/R}, \underline{\omega}^k_{/R})$ with $k \in \mathbb{Z}[I]$. We have the following assertions.*

(1) $f = 0 \iff a_{\mathfrak{a},\mathfrak{b}}(\xi, f) = 0$ *for all* $\xi \in (\mathfrak{a}\mathfrak{b})_+ \cup \{0\}$;
(2) $f \in H^0(T_{\alpha/R'}, \underline{\omega}^k_{/R'}) \iff a_{\mathfrak{a},\mathfrak{b}}(\xi, f) \in R'$ *for all* $\xi \in (\mathfrak{a}\mathfrak{b})_+ \cup \{0\}$.

One may state a similar assertion for p-adic modular forms, which is left for the reader to formulate.

Proof. The first assertion follows from the irreducibility of the scheme T_α. The same assertion holds for any R-module M and $H^0(T_{\alpha/R}, \underline{\omega}^k_{/R} \otimes_R M)$ for the same reason. To see (2), we look at the exact sequence

$$0 \to H^0(T_{\alpha/R'}, \underline{\omega}^k_{/R'}) \to H^0(T_{\alpha/R}, \underline{\omega}^k_{/R}) \to H^0(T_{\alpha/R'}, \underline{\omega}^k_{/R'} \otimes (R/R')).$$

Note that $f \in H^0(T_{\alpha/R'}, \underline{\omega}^k_{/R'})$ if and only if the image of f vanishes in $H^0(T_{\alpha/R'}, \underline{\omega}^k_{/R'} \otimes (R/R'))$. Applying the module version of (1) for $M = R/R'$, we get (2). □

Let $R = \varprojlim_m R/p^mR$ be a p-adically complete W-algebra. We write Ig for the Igusa tower $T_{\infty,\infty}$ over $S_\infty = \varprojlim_m Sh^{(p)}\left[\frac{1}{E}\right] \otimes_W R/p^mR$ for a lift $E \in H^0(S_\infty, \underline{\omega}^{a(p-1)I})$ of the Hasse invariant H. For a subgroup K with $S_1(p^r\mathfrak{N}) \subset K \subset S_0(p^r\mathfrak{N})$ for an integral ideal $\mathfrak{N}$ prime to p, the group $K/S_1(p^r\mathfrak{N})$ can be identified with a subgroup of $T(\mathbb{Z}/p^r\mathbb{Z})$. We write $K'_p \subset T(\mathbb{Z}_p)$ for the pullback of this subgroup $K/S_1(p^r\mathfrak{N}) \subset T(\mathbb{Z}/p^r\mathbb{Z})$. Then by abusing the notation, we

write Ig_K for the quotient $Ig/(K'_p \times K^{(p)})$, where K'_p acts on Ig through the identification of $T(\mathbb{Z}_p)$ with $\mathrm{Gal}(Ig/Sh^{(p)})$. This is consistent because Ig_K gives an open irreducible subscheme of the reduction modulo p of the integral model over W of Sh_K relative to the level K-structure of Drinfeld type (see [AME] Chapter 3, [GME] 2.6.2 and 2.9, and [GCS] II.2). Thus we have $Ig_{S_0(p^r\mathfrak{N})} = S_\infty$ which is independent of r (but $Ig_{S_0(p^r\mathfrak{N})} \subsetneq Sh^{(p)}_{S_0(\mathfrak{N})}$).

4.2.5 Hecke Operators as Algebraic Correspondences

We write Sh_K for $Sh_K(G, X)$. For any $g \in G(\mathbb{A}^{(\infty)})$ and open compact subgroups $K, K' \subset G(\mathbb{A}^{(\infty)})$, we have two projections $p_1 : Sh_{K^g \cap K'} \to Sh_{K'}$ given by $([x,h] \mod K^g \cap K') \mapsto ([x,h] \mod K')$ and $p_g : Sh_{K^g \cap K'} \to Sh_K$ given by $([x,h] \mod K^g \cap K') \mapsto ([x, hg^{-1}] \mod K)$, where $K^g = g^{-1}Kg$ and ${}^gK' = gK'g^{-1}$. The image of $p_g \times p_1$ in $Sh_K \times_{\mathbb{Q}} Sh_{K'}$ is an algebraic correspondence $(KgK') \subset Sh_K \times_{\mathbb{Q}} Sh_{K'}$ which only depends on the double coset KgK' as easily verified, using the fact $Sh_K = Sh/K$. We now assume

(s1) We have a vector bundle $\mathcal{L}_{/Sh}$ (either an algebrogeometric locally free sheaf of finite rank on the algebraic variety $Sh_{/\mathbb{Q}}$ or a locally constant sheaf on the topological space $Sh(\mathbb{C})$) which descends to a sheaf $\mathcal{L}_{K/Sh_K}$ of the same type for sufficiently small K so that $\pi^*\mathcal{L}_K = \mathcal{L}$ for the projection $\pi : Sh \to Sh_K$;

(s2) There exists an open semi-group $\Delta \subset G(\mathbb{A}^{(\infty)})$ such that Δ acts on $\mathcal{L}_{/Sh}$ via g^* making the following diagram commutative,

$$\begin{array}{ccc} \mathcal{L} & \xleftarrow{g^*} & \mathcal{L} \\ \downarrow & & \downarrow \\ Sh & \xrightarrow[g^{-1}]{} & Sh, \end{array}$$

where the bottom arrow is given by the action of $g^{-1} \in \Delta^{-1}$ on Sh and the top arrow is the pullback left action (because the action of $G(\mathbb{A}^{(\infty)})$ on Sh is a right action). For an open compact subgroup $K \subset \Delta$, we require that the quotient $\mathcal{L}/K$ under the above action k^* of $k \in K$ give rise to $\mathcal{L}_{K/Sh_K}$; so, we have $H^0(Sh_K, \mathcal{L}_K) = H^0(K, H^0(Sh, \mathcal{L}))$;

(s3) We have the trace map on the cohomology groups

$$\mathrm{Tr}_{K'/K} : H^\bullet(Sh_K, \mathcal{L}_K) \to H^\bullet(Sh_{K'}, \mathcal{L}_{K'})$$

for $K \subset K' \subset \Delta$ satisfying that $\mathrm{Tr}_{K''/K'} \circ \mathrm{Tr}_{K'/K} = \mathrm{Tr}_{K''/K}$ for $K \subset K' \subset K''$, and $\mathrm{Tr}_{K'/K} \circ \mathrm{Res}_{K'/K}$ for the restriction map $\mathrm{Res}_{K'/K}$ is the multiplication by the covering degree of $Sh_K \to Sh_{K'}$.

Since Sh/Sh_S is étale and Sh_S is smooth if S is sufficiently small, choosing such an open compact subgroup S and a sheaf $\mathcal{L}_S$ on Sh_S as in (s1), the

two conditions (s1) and (s3) are satisfied by the quotient $\mathcal{L}/K$ for $\mathcal{L} = \pi^*\mathcal{L}_S$ as long as $K \subset S$. As we show in the following discussion, the invertible sheaf $\underline{\omega}_K^k = \underline{\omega}^k_{/Sh_K}$ and the locally constant sheaf $L(\kappa; R)_K = L(\kappa; R)_{/Sh_K}$ presented in the introduction satisfy these properties.

Hereafter in this section, we limit ourselves to coherent sheaves $\mathcal{L}_K$ (topological sheaves are dealt with in Section 4.3.3). Under (s1–2), for subgroups K and K' of Δ, we define the *Hecke operator* $[KgK'] : H^\bullet(Sh_K, \mathcal{L}_K) \to H^\bullet(Sh_{K'}, \mathcal{L}_{K'})$ associated with the correspondence (KgK') by

$$[KgK'] = |\det(g)|_{\mathbb{A}} \cdot \mathrm{Tr}_{K'/(K^g \cap K')} \circ [g] \circ \mathrm{Res}_{K/(K \cap {}^gK')}, \tag{4.37}$$

where $[g] : H^\bullet(S_{K \cap {}^gK'}, \mathcal{L}_{K \cap {}^gK'}) \to H^\bullet(Sh_{K^g \cap K'}, \mathcal{L}_{K^g \cap K'})$ is induced by the pullback sheaf morphism g^* of the left action $g^{-1} : Sh_{K^g \cap K'} \cong Sh_{K \cap {}^gK'}$, and $|a|_{\mathbb{A}}$ is the adele norm of $a \in F_{\mathbb{A}}^\times$.

In this way, for any open compact subgroup K of $G(\mathbb{A}^{(\infty)})$ inside Δ, the double coset ring $R(K, \Delta)$ with respect to $K \subset \Delta$ acts on the cohomology group $H^\bullet(Sh_K, \mathcal{L}_K)$. Here $R(K, \Delta)$ is the free $\mathbb{Z}$-module formally generated by double cosets KgK for $g \in \Delta$ with the multiplication law as given in [IAT] 3.1 (or [MFG] 3.1.6; see also Section 5.1.2 in the text). We call $R(K, \Delta)$ the *double coset ring* for $K \subset \Delta$. An important point is that we need to specify the action g^* of $g \in \Delta$, and a given sheaf $\mathcal{L}$ may have different actions (which could be canonical from different points of view). The module structure of the cohomology group over $R(K, \Delta)$ depends on the action of Δ on $\mathcal{L}_K$.

For subgroups K with $S_1^1(\mathfrak{N}) \subset K \subset S_0(\mathfrak{N})$, we define a standard semi-group $\Delta_0(\mathfrak{N})$ as follows. Let $\Delta^{(\mathfrak{N})} = M_2(\widehat{O}) \cap GL_2(F_{\mathbb{A}^{(\mathfrak{N}\infty)}})$. At $\mathfrak{N}$, we define

$$D = \left\{ \left(\begin{smallmatrix} a & 0 \\ 0 & d \end{smallmatrix}\right) \in M_2(O_{\mathfrak{N}}) \middle| a \in O_{\mathfrak{N}}^\times,\ d \in O_{\mathfrak{N}} \right\},$$

and put $\Delta_{\mathfrak{N}} = (S_0(\mathfrak{N})D \cdot S_0(\mathfrak{N})) \cap GL_2(F_{\mathfrak{N}})$. These semi-groups are characterized by the fact that $\delta \cdot U(O_{\mathfrak{N}})\delta^{-1} \supset U(O_{\mathfrak{N}})$ for the upper triangular unipotent subgroup $U \subset G$ (so D and $\Delta_{\mathfrak{N}}$ are called *expanding* semi-groups at $\mathfrak{N}$). The involution $g \mapsto \det(g)g^{-1} = g^\iota$ brings an expanding semi-group to the opposite shrinking one. We define $\Delta_0(\mathfrak{N}) = \Delta_{\mathfrak{N}} \times \Delta^{(\mathfrak{N})}$.

4.2.6 Modular Line Bundles

The diagonal torus T_G of G is isomorphic to T^2; so, its weight κ is a pair (κ_1, κ_2) of weights of T. We associate with each dominant weight κ and a "Neben" character ε modulo $\mathfrak{N}$, a coherent line bundle $\underline{\omega}_\varepsilon^\kappa$ and its cuspidal sub-bundle $\underline{\omega}_{\kappa,\varepsilon}$ equipped with the action of $\Delta_0(\mathfrak{N})$ satisfying the three conditions (s1–3) and giving rise to the space of modular forms or cusp forms of weight κ and character ε. Since we have well-defined $\underline{\omega}^k$ and $\underline{\omega}_k$ for each dominant weight k of T on each geometrically connected component of Sh or $Sh^{(p)}$, we have well-defined $\underline{\omega}^k$ and $\underline{\omega}_k$ on Sh. Thus our discussion is focused on how to define a good action of $\Delta_0(\mathfrak{N})$ on these line bundles. The quotient $\underline{\omega}_K^k$ over Sh_K equipped with thus-defined action g^* of $g \in \Delta$ really depends on

the weight κ of T_G (not just the weight k of T) and the character ε. This involves a fair amount of technicality. However, once we succeed, we will have a coherent line bundle with eigensections (of all Hecke operators) associated with motives of the Hodge type directly related to the weight κ, whose Galois representations have determinant character dependent only on ε and κ. Thus the reward for our endeavor is worth going through this technicality. We fix a base ring $\mathcal{B}$ which is often one of the following rings: the valuation ring $\mathcal{W}$ of residual characteristic $p > 0$, its residue field $\mathbb{F}$, $\mathbb{F}_p$, $\mathbb{Z}_{(p)}$, $\mathbb{Q}$, or $\mathbb{C}$ (in any case, $\mathcal{B}$ is an algebra over the ones specified above). Thus the line bundle and its section are supposed to be defined over a $\mathcal{B}$-algebra R.

We now start our work. We have a natural map $\mathfrak{M}'(\mathfrak{c}, \Gamma(N)) \to Sh_{S(N)}$ bringing (A, λ, ϕ_N) to $(A, \overline{\lambda}, \overline{\eta})$ so that $\eta \mod S(N)$ gives rise to ϕ_N. On $\mathfrak{M}'(\mathfrak{c}, \Gamma(N))$, we have a well-defined invertible sheaf $\underline{\omega}^k$. Since $Sh_{S(N)}$ is the union of its geometrically irreducible components $\mathfrak{M}'(\mathfrak{c}, \Gamma(N))$, we have well-defined $\underline{\omega}^k_{S(N)}$ for $N \geq 3$ (and its cuspidal counterpart $\underline{\omega}_{k,S(N)}$) on $Sh_{S(N)}$, which after taking pullback to Sh, gives rise to a line bundle $\underline{\omega}^k$ defined over Sh. This sheaf satisfies (s1) and (s3), but the resulting sheaf $\underline{\omega}^k$ and the Hecke operators acting on it would depend on our choice of the action of Δ; so, we need to clarify this to have a clear picture of the modular forms on Sh and the Hecke operator action.

Let us explain this process of defining $\underline{\omega}^k$ more carefully. We first choose a complete representative set $\{\mathfrak{c}\}$ of strict ideal classes of F. We choose $\mathfrak{c}$ prime to p. Once $\mathfrak{c}$ is chosen, we can always specify a $\mathfrak{c}$-polarization λ in the polarization class $\overline{\lambda}$. The endomorphism $\iota(a) \in \mathrm{End}(A)$ $(a \in F)$ acts on λ by $\iota(a)^*\lambda = a^2\lambda$ for a test object $(A, \lambda, \eta^{(p)})$. Thus even if we take a $\mathfrak{c}$-polarization λ, for a unit $\epsilon \in O^\times$, $(A, \iota(\epsilon)^*\lambda, \eta^{(p)} \circ \epsilon)$ is another test object with $\mathfrak{c}$-polarization representing the same point of $Sh^{(p)}$.

The problem in defining the action g^* on $\underline{\omega}^k_{/Sh}$ is that each point of $Sh^{(p)}_K$ corresponds to an isogeny class (not a single abelian variety); so, we have inherent ambiguity coming from suppleness of isogeny. In other words, the point here is to associate without ambiguity a section ω^k of $\underline{\omega}^k$ with each section ω of $H^0(A, \Omega_{A/S})$ for a given $(A, \lambda, \eta^{(p)}) \in \mathcal{E}^{(p)}_K(S)$ only depending on the point of Sh representing the test object.

We can do this job of specifying ω^k functorially by (connected) component by component; so, we may assume that $\overline{\lambda}$ determines a strict ideal class of $\mathfrak{c}$. Then we choose a representative λ with the given polarization ideal $\mathfrak{c}$. Let $L = \mathfrak{d}^{-1} \oplus \mathfrak{c}^{-1} \subset V$ and $\Lambda : L \wedge_O L \cong \mathfrak{c}^*$ be induced by the pairing $\langle (a, a'), (b, b') \rangle = a'b - ab'$. We have a canonical alternating pairing $\langle \cdot, \cdot \rangle_\lambda : V^{(p)}(A) \times V^{(p)}({}^tA) \to F_{\mathbb{A}^{(p\infty)}}$, which via λ and the level structure η induces a linear map $\Lambda^{(p)} : \widehat{L}^{(p)} \wedge_{\widehat{O}^{(p)}} \widehat{L}^{(p)} \cong (\widehat{\mathfrak{c}}^*)^{(p)}$. We consider a rigidity condition:

$$\Lambda^{(p)} = \lambda^{(p)}. \tag{4.38}$$

If we change η by $\eta \circ u$ for $u \in K^{(p)}$, this condition fails if $\det(u) \neq 1$. Thus under (4.38), we need to assume that $K^{(p)} \subset G_1(\mathbb{A}^{(p\infty)})$ in order to make the class $\overline{\eta} = \eta K^{(p)}$ compatible with this condition.

For a finite étale morphism $A_{/S} \xrightarrow{\alpha} Y_{/S}$, we have a canonical exact sequence (e.g., [GME] Proposition 1.5.4) $\Omega_{Y/S} \xrightarrow{\alpha^*} \Omega_{A/S} \to \Omega_{A/Y} \to 0$. If α is étale, α^* is an isomorphism; so, we have an identification $\alpha^* : \Omega_{Y/S} \cong \Omega_{A/S}$. If $\alpha : A \to Y$ is an étale isogeny, we can factor $\alpha = \beta \circ \gamma^{-1}$ for finite étale morphisms $A' \xrightarrow{\gamma} A$ and $A' \xrightarrow{\beta} Y$. In this case, we define $\alpha^* : H^0(Y, \Omega_{Y/S}) \cong H^0(A, \Omega_{A/S})$ by $\alpha^* = (\gamma^*)^{-1} \circ \beta^*$.

For $(A, \lambda, \overline{\eta}^{(p)})_{/S} \in \mathcal{E}_K^{(p)}(S)$ with $\mathfrak{c}$-polarization λ, as was done in Section 4.2.1, bringing the functor $\mathcal{E}_K^{(p)}$ to $\mathcal{E}'_K$ isomorphically, we find an étale isogeny $\alpha : (A, \lambda, \eta^{(p)}) \to (Y, \lambda', \eta'^{(p)})$ over S with $\eta'^{(p)}(\widehat{L}) = \mathcal{T}(Y) \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}^{(p)}$ and a $\mathfrak{c}$-polarization λ'. For the point $x \in Sh^{(p)}_{S(N)}(S)$ representing $(A, \lambda, \eta^{(p)})$ and $\omega \in H^0(A, \Omega_{A/S})$, our idea is to take $(\alpha^{-*}\omega)^{\otimes k}$ as a section (representing $\omega^{\otimes k}$) of $\underline{\omega}^k_{S(N)}$ at x. Here we have written α^{-*} for $(\alpha^*)^{-1}$ for simplicity. This idea works well under (4.38). Indeed, if we have another isogeny $\beta : (A, \lambda, \eta^{(p)}) \to (Y, \lambda', \eta'^{(p)})$ over S with $\eta'^{(p)}(\widehat{L}) = \mathcal{T}(Y) \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}^{(p)}$, the composite $\alpha^{-1} \circ \beta$ is an automorphism of $(Y, \lambda', \eta'^{(p)})_{/S} \in \mathcal{E}'_K(S)$, which has to be induced by a unit $\epsilon \in O^\times$ with $\epsilon \in K$ and $\epsilon^* \lambda' = \lambda'$ by (4.38). It is easy to see that if $K^{(p)} \subset S(N)^{(p)}$ with $p \nmid N \geq 3$, the unit ϵ has to be trivial. Thus we have α without ambiguity, and hence $(\alpha^{-*}\omega)^{\otimes k}$ without ambiguity.

Without assuming (4.38), just to prove the representability of $\mathcal{E}_K^{(p)}$, bringing it isomorphically to $\mathcal{E}'_K$ by an isogeny normalized with respect to the lattice L works perfectly well, but this process lacks rigidity to determine the section $\omega^{\otimes k}$ of $\underline{\omega}^k$ uniquely. Since the central element $\xi \in Z(\mathbb{Q})$ acts through the endomorphism $\iota(\xi) \in \mathrm{End}_F^{\mathbb{Q}}(A)$ on $\omega \in H^0(A_{/\mathbb{C}}, \Omega_{A/\mathbb{C}})$ by $\iota(\xi)^*\omega = \xi_\infty \omega$, we enlarge our group $G(\mathbb{Z}_p \times \mathbb{A}^{(p\infty)})$ to $G(\mathbb{Z}_p \times \mathbb{A}^{(p)})^+ := G(\mathbb{Z}_p) \times G(\mathbb{A}^{(p\infty)}) \times G(\mathbb{R})^+$ and let $G(\mathbb{Z}_p \times \mathbb{A}^{(p)})^+$ act on $Sh^{(p)}$ through its quotient $G(\mathbb{Z}_p \times \mathbb{A}^{(p)})^+ / G(\mathbb{R})^+$. We write $Sh^{(p)}_K$ for $Sh^{(p)}/K^{(\infty)}$ for a closed subgroup K in $G(\mathbb{A})^+$ with $K_\infty = G(\mathbb{R})^+$ and $K_p = G(\mathbb{Z}_p)$. Thus we have eased the condition $K \subset G(\mathbb{Z}_p) \times G_1(\mathbb{A}^{(p\infty)})$ to a milder one: $K \subset G(\mathbb{Z}_p \times \mathbb{A}^{(p)})^+$. Each element in $K \cap Z(\mathbb{Q})$ is a root of unity up to units in O; so, if $k = mI$ for an integer m (and K is sufficiently small), $(\alpha^{-*}\omega)^{\otimes mI}$ is uniquely determined, since $\epsilon^{2mI} = 1$ for all $\epsilon \in O^\times$. Our argument now gives a well-defined $\underline{\omega}^{mI}_K$ for a sufficiently small open compact subgroup $K \subset G(\mathbb{Z}_p \times \mathbb{A}^{(p)})^+$ (maximal at p) and a well-defined $\underline{\omega}^k_K$ for general k with a sufficiently small open compact subgroup $K \subset G(\mathbb{Z}_p) \times G_1(\mathbb{A}^{(p\infty)}) \times G(\mathbb{R})^+$.

We write $K_1 = (K \cap G_1(\mathbb{A}^{(p\infty)})) \times G(\mathbb{R})^+$ for a given K in $G(\mathbb{Z}_p \times \mathbb{A}^{(p)})^+$. We have a well-defined $\underline{\omega}^k$ over $Sh^{(p)}_{K_1} = Sh^{(p)}/K_1^{(\infty)}$ for a sufficiently small K; so, we have $\underline{\omega}^k$ over $Sh^{(p)}$ just by pullback. If we can define a good action of $G(\mathbb{Z}_p \times \mathbb{A}^{(p)})^+$ on $\underline{\omega}^k$ over $Sh^{(p)}$, we can descend $\underline{\omega}^k_{K_1}$ to $\underline{\omega}^k_K$ so that the

global sections over $Sh_K^{(p)}$ are made up of global sections over $Sh_{K_1}^{(p)}$ fixed by K. Similarly, if we can define a good action of $T_G(\mathbb{Z}_p) \times G(\mathbb{A}^{(p)})^+$ on $\underline{\omega}^k$ over the Igusa tower $Ig_{/Sh^{(p)}}$, we can descend $\underline{\omega}^k_{/Ig}$ to $\underline{\omega}^k_K$ for K with $S_1(p^n)_p \subset K_p \subset G(\mathbb{Z}_p)$.

For the moment, we assume that the base ring $\mathcal{B}$ is $\mathcal{W}$. To define the desired $\mathcal{W}$-integral action of Δ, starting with a K_1-test object $(A, \lambda, \overline{\eta}^{(p)} \circ g^{-1})$, we define $(Y_g, \lambda_g, \overline{\eta}_g{}^{(p)}) \in \mathcal{E}'_{{}^gK_1}$ for ${}^gK_1 = gK_1g^{-1}$ equivalent to $(A, \lambda, \eta^{(p)} \circ g^{-1})$ uniquely in the same manner as above, replacing $\eta^{(p)}$ by $\eta^{(p)} \circ g^{-1}$, since we have the left (inverted) action g^{-1} of g in (s2). We need to assume that g^{-1} for $g \in \Delta$ has an associated étale isogeny $\alpha_g : (A, \lambda, \eta^{(p)} \circ g^{-1}) \to (Y_g, \lambda_g, \eta_g^{(p)})$ for all p-ordinary test objects $\underline{A} = (A, \lambda, \eta^{(p)})$. Here we say $\underline{A}$ p-ordinary if A is ordinary at p. Also we mean by an "étale isogeny" $\alpha_g : A \to Y_g$ an isogeny given by $\beta \circ \alpha^{-1}$ for étale finite morphisms $A' \xrightarrow{\alpha} A$ and $A' \xrightarrow{\beta} Y_g$. Let $\Delta_0(p^\infty\mathfrak{N}) = \bigcap_r \Delta_0(p^r\mathfrak{N})$. If $g \in \Delta_0(p^\infty\mathfrak{N})$, then α_g is an étale isogeny. Thus we assume that Δ^{-1} is contained in

$$\left\{ g \in G(\mathbb{A})^+ \middle| \alpha_g \text{ is étale for all } p\text{-ordinary } (A, \lambda, i_p, \eta^{(p)})_{/S} \right\}, \tag{4.39}$$

where $i_p : \mu_p \otimes \mathfrak{d}^{-1} \hookrightarrow A[p]$ is a p-ordinary level structure, and S runs over all $\mathcal{W}$-schemes. Strictly speaking, the isogeny α_g can be defined over a faithfully flat covering S'/S, which could be inseparable (see Section 4.4.2 for separability of a morphism of schemes).

$$\text{If } \mathcal{B} = \mathcal{W}, \text{ we require } \Delta_p = \Delta_0(p^\infty)_p. \tag{4.40}$$

If $\mathcal{B} = \mathcal{W}$, we may extend i_p to a level p^∞-structure $\eta_p^{ord} : \mu_{p^\infty} \otimes \mathfrak{d}^{-1} \hookrightarrow A_{/S'}$ for the base change $A_{/S'}$ of $A_{/S}$ to an étale covering S'/S. Normalize the first coordinate a of $\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) \in \mathfrak{a}^* \oplus \mathfrak{b} = L$ ($a \in \mathfrak{a}^*$ and $b \in \mathfrak{b}$) so that it is associated with η_p^{ord}. Then the above choice (4.40) of Δ satisfies the requirement (4.39) for $\underline{A}$, because $g_p \in \Delta_p$ by definition preserves the a coordinate and hence the connected subgroup $i_p(\mu_p \otimes \mathfrak{d}^{-1})$ inside $A[p]$. The normalization of coordinates is equivalent to choosing a point carrying $(A, \eta_p^{ord}, \eta^{(p)})$ on the component of the Hilbert modular Igusa tower containing the cusp $(\mathfrak{a}, \mathfrak{b})$ (see the proof of Theorem 4.21). Hereafter we write η for $\eta_p^{ord} \times \eta^{(p)}$ in order to treat the case of $Sh^{(p)}$ and Ig uniformly. When we are dealing with $Sh^{(p)}$, we just ignore η_p^{ord} and the ordinarity of A.

We may further impose (4.38). If we do impose (4.38), the isogeny

$$\alpha_g : (A, \overline{\lambda}, \overline{\eta} \circ g^{-1}) \to (Y_g, \overline{\lambda}_g, \overline{\eta}_g) \quad \text{with} \quad \alpha_g \circ \overline{\eta} \circ g^{-1} = \overline{\eta}_g$$

uniquely determines $g^{-1}\omega = \alpha_g^{-*}\omega$ for $\omega \in H^0(A, \Omega_{A/S})$ (which is a section of $\underline{\omega}_{gK}$ over Sh_{gK}). Since $\omega^{\otimes k}$ is a section of $\underline{\omega}^k_K$ over Sh_K at $(A, \overline{\lambda}, \overline{\eta})$, we have a natural left action $\omega^{\otimes k} \mapsto g^{-1} \cdot \omega^{\otimes k} := (g^{-1}\omega)^{\otimes k} = (\alpha_g^{-*}\omega)^{\otimes k}$ of $g \in \Delta$ making the following diagram commutative:

$$\begin{array}{ccc} \underline{\omega}^k_{K_1} & \xrightarrow{g^{-1}} & \underline{\omega}^k_{g\,K_1} \\ \downarrow & & \downarrow \\ Ig_{K_1} & \xrightarrow[g^{-1}]{} & Ig_{g\,K_1}\,, \end{array}$$

where the bottom arrow is given by $[x,h] \mapsto [x, hg^{-1}]$.

If $\mathcal{B}$ is a $\mathbb{Q}$-algebra, we can ignore the condition related to $v = p$ (in particular i_p) in (4.39), because ordinarity and étaleness in (4.39) are always satisfied in this case. Taking now $\mathcal{B} = \mathbb{Q}$ and $\Delta = G(\mathbb{A})^+$, by the same process as above, we can define the action of $G(\mathbb{A})^+$ on $\underline{\omega}^k$ over the Shimura variety $Sh_{/\mathbb{Q}}$. If $\mathcal{B}$ is a $\mathcal{W}$-algebra, similarly, we have a well-defined left action of $G(\mathbb{A}^{(p\infty)})$ on differentials over $Sh^{(p)}_{/\mathcal{W}}$ (and $Ig_{/\mathcal{W}}$).

The problem we now encounter is that the action of Δ we have so far defined may not factor through $G(\mathbb{A})^+/\overline{Z(\mathbb{Q})}$. In other words, the descent to Sh_K we desired to have is for an open compact subgroup K of $G(\mathbb{Z}_p \times \mathbb{A}^{(p)})^+$, but so far, we described it for K inside $G(\mathbb{Z}_p) \times G_1(\mathbb{A}^{(p\infty)}) \times G(\mathbb{R})^+$. To make this wider descent allowing all $K^{(p)}$ in $G(\mathbb{A}^{(p)})^+$, we want to have the action factoring through $K/(K \cap \overline{Z(\mathbb{Q})})$ in order to make the quotient of the sheaf over Sh by K. We have an isogeny $\epsilon : (A, \epsilon^{-2}\lambda, \eta, \epsilon\omega) \to (A, \lambda, \eta \circ \epsilon, \omega)$ for $\epsilon \in O^\times$; so, $\epsilon \cdot \omega = \epsilon\omega$. Thus the problem is to somehow modify the unit action on differentials $\omega^{\otimes k} \mapsto \epsilon^k \omega^{\otimes k}$ for $\epsilon \in K \cap Z(\mathbb{Q}) \subset O^\times$. This problem can be solved if the base ring is complete; so, we assume that $\mathcal{B} = \mathbb{C}$ or $W = \varprojlim_n \mathcal{W}/p^n\mathcal{W}$. We write v for the place p or ∞ (of $\mathbb{Q}$), and use the symbol $\mathcal{B}$ to indicate the base ring depending on our choice of v; so, $\mathcal{B} = \mathbb{C}$ if $v = \infty$ and $\mathcal{B} = W$ if $v = p$.

To state the solution (found first in [H91] Section 3 and later developed and brought into the present form in Fujiwara [F1]), recall the diagonal torus T_G of G; so, $T_G \cong T^2$. We choose a character $w \in \mathbb{Z}[I] = X(T)$ and a continuous character $\varepsilon : T_G(\widehat{\mathbb{Z}}) \to \mathcal{W}^\times$ such that $\varepsilon_+(\epsilon)^{-1}\epsilon^{k-2w} = 1$ for all units ϵ, where $\varepsilon_+ : Z(\widehat{\mathbb{Z}}) \to \mathcal{W}^\times$ (for the center Z of G) is given by $\varepsilon_+(z) = \varepsilon_1(z)\varepsilon_2(z)$ writing $\varepsilon(\mathrm{diag}[a,d]) = \varepsilon_1(a)\varepsilon_2(d)$. Define $\varepsilon^- : T(\widehat{\mathbb{Z}}) \to \mathcal{W}^\times$ by $\varepsilon^-(z) = \varepsilon_2^{-1}(z)\varepsilon_1(z)$. We assume that ε_+ extends to a Hecke character $\varepsilon_+ : Z(\mathbb{A})/Z(\mathbb{Q}) \to \mathbb{C}^\times$ with $\varepsilon_+(x_\infty) = x_\infty^{2w-k}$. Here $\mathcal{W}$ is equipped with the discrete topology; so, $\varepsilon_\jmath$ is of finite order, although ε_+ extended to $Z(\mathbb{A})$ could be of infinite order. In particular, ε^- factors through $(O/\mathfrak{N})^\times$ for an ideal $\mathfrak{N}$.

We fix a prime element $\varpi_\mathfrak{q}$ for each prime ideal $\mathfrak{q}$. We suppose that $\varpi_\mathfrak{p} = p$ for $\mathfrak{p}|p$ (because p is unramified in $F/\mathbb{Q}$). Sending $x \in F_\mathfrak{q}^\times$ to $(u, \varpi_\mathfrak{q}^e) \in O_\mathfrak{q}^\times \times \varpi_\mathfrak{q}^\mathbb{Z}$ if $x = u\varpi_\mathfrak{q}^e$, we have the identification $F_\mathfrak{q}^\times = O_\mathfrak{q}^\times \times \varpi_\mathfrak{q}^\mathbb{Z}$. Hence $T(\mathbb{A}^{(\infty)}) = T(\widehat{\mathbb{Z}}) \times \left(\bigoplus_\mathfrak{q} \varpi_\mathfrak{q}^\mathbb{Z}\right)$. Recall that $v = p$ or ∞ is the fixed place of $\mathbb{Q}$. We now extend our Neben character $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_+)$ to $T(\mathbb{A}^{(\infty)})$ as follows.

(ex0) *For each $w \in \mathbb{Z}[I]$, we extend the character $x_p \mapsto x_p^w$ of $T(\mathbb{Z}_p)$ to $T(\mathbb{Q}_p)$ trivially on $\bigoplus_{\mathfrak{p}|p} \varpi_\mathfrak{p}^\mathbb{Z}$. Thus $p^{-2I} = 1$.*

(ex1) *We extend* ε_2 *to* $T(\mathbb{A}^{(\infty)})$ *trivially on* $\bigoplus_{\mathfrak{q}} \varpi_{\mathfrak{q}}^{\mathbb{Z}}$ *and then extend* ε_1 *to* $T(\mathbb{A}^{(\infty)})$ *by* $\varepsilon_1\varepsilon_2(x) = \varepsilon_+(x^{(\infty)})$. *We put* $\varepsilon_v^-(a) = \varepsilon_2^{-1}(a)\varepsilon_1(a)$ *for* $a \in T(\mathbb{A}^{(\infty)})$. *Thus identifying* $T^2 \cong T_G$ *by* $(a,d) \mapsto \left(\begin{smallmatrix} a & 0 \\ 0 & d \end{smallmatrix}\right)$, *we get an extension of* ε *to* $T_G(\mathbb{A})$.

(ex2) *We extend the character* ε *of* $T_G(\widehat{\mathbb{Z}}) \cap S_0((v) \cap \mathfrak{N})$ *to a character of the semi-group* $\Delta_0((v) \cap \mathfrak{N})$ *by*

$$\varepsilon_\Delta^v(\delta) = \det(\delta)_v^{-I}\varepsilon_2(\det(\delta))\varepsilon_v^-(a_{(v)\cap\mathfrak{N}})$$

for $\delta = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Delta_0((v) \cap \mathfrak{N}) \cup \Delta_0((v) \cap \mathfrak{N})^\iota$. *Here* $(v) \cap \mathfrak{N} = (p) \cap \mathfrak{N}$ *if* $v = p$ *and* $(v) \cap \mathfrak{N} = \mathfrak{N}$ *if* $v = \infty$.

(ex3) *Since the character* $Z(\mathbb{A}^{(\infty)})/Z(\mathbb{Q}) \ni z \mapsto |z^{(\infty)}|_{\mathbb{A}}^{-2} z_v^{-2I}\varepsilon_+(z)$ *and* ε_Δ^v *on* $S_0((v) \cap \mathfrak{N})$ *coincide on* $S_0((v) \cap \mathfrak{N}) \cap Z(\mathbb{A}^{(\infty)})$, *we may extend* $\varepsilon_\Delta^v|_{S_0((v)\cap\mathfrak{N})}$ *to* $\varepsilon_S = \varepsilon_S^v : Z(\mathbb{A})S_0((v)\cap\mathfrak{N}) = Z(\mathbb{A})S_1((v)\cap\mathfrak{N}) \to \mathcal{B}^\times$ *by* $\varepsilon_S^v(zs) = |z^{(\infty)}|_{\mathbb{A}}^{-2} z_v^{-2I}\varepsilon_+(z)\varepsilon_\Delta^v(s)$ *for* $z \in Z(\mathbb{A})$ *and* $s \in S_1((v) \cap \mathfrak{N})$.

If we need to emphasize the dependence on v, we add the superscript v, as in ε_S^v for ε_S. Note here that ε_S depends on v but the central character ε_+ and ε_j $(j = 1, 2)$ do not. Also $\varepsilon_1(\varpi_{\mathfrak{q}})$ could be nontrivial, but $\varepsilon_2(\varpi_{\mathfrak{q}}) = 1$.

By computation, we have

$$\varepsilon_S^v(z) = \varepsilon_+(z)|z^{(\infty)}|_{\mathbb{A}}^{-2} z_v^{-2I} \quad \text{for } z \in Z(\mathbb{A}),$$
$$\text{and } \varepsilon_\Delta^v\left(\left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right)_{(v)\cap\mathfrak{N}}\right) = a_v^{-I} d_v^{-I}\varepsilon_1(a_{(v)\cap\mathfrak{N}})\varepsilon_2(d_{(v)\cap\mathfrak{N}}). \quad (4.41)$$

We define new actions $\omega^{\otimes k} \mapsto g_{w,\varepsilon,v}\omega^{\otimes k}$ of $g \in Z(\mathbb{A})S_0(\mathfrak{N})$ and $\omega^{\otimes k} \mapsto g^*\omega^{\otimes k}$ in (s2) of $g \in \Delta_0(\mathfrak{N})$ (or $g \in \Delta_0(\mathfrak{N})^\iota$) by

$$g_{w,\varepsilon,v}\omega^{\otimes k} = \varepsilon_S^v(g)^{-1}\det(g)_v^{-w}(g \cdot \omega^{\otimes k})$$
$$\text{and } g^*(\omega^{\otimes k}) = \varepsilon_\Delta^v(g)\det(g)_v^w(g^{-1} \cdot \omega^{\otimes k}). \quad (4.42)$$

The first action is a left action, and the second is considered to be a right pullback action associated with $Sh \ni x \mapsto xg^{-1} \in Sh$ for $g \in \Delta_0(\mathfrak{N})$. These two actions coincide on $S_0(\mathfrak{N}) \cap Z(\mathbb{Q}) = Z(\mathbb{Z})$ and commute each other, and hence the action of $[KgK']$ with respect to the second action is well-defined on the cohomology group of the coherent sheaf on Sh_K obtained from $\underline{\omega}^k_{/Sh}$ by the descent via the first action. For the second action to be valid, one has to assume that α_g is an étale isogeny (which always holds if $v = \infty$). Then, writing $\xi^{(v\infty)}$ for $\xi^{(\infty)}$ or $\xi^{(p\infty)}$ accordingly as $v = \infty$ or p, we have $\xi_{w,\varepsilon,v}\omega^{\otimes k} = |\xi^{(v\infty)}|_{\mathbb{A}}^{-2}\xi_v^{-2I}\varepsilon_+(\xi^{(\infty)})^{-1}\xi^{-2w}\xi^k\omega^{\otimes k} = \omega^{\otimes k}$ (for $\xi \in O_{(v)+}^\times$) by the product formula $|\xi^{(v\infty)}|_{\mathbb{A}}^2\xi_v^{2I} = 1$, and we are done. Here $O_{(v)+}^\times = O_{(p)+}^\times$ if $v = p$ and $O_{(v)+}^\times = F_+^\times$ if $v = \infty$. The action of $K^{(p)}$ is well-defined over $\mathcal{W}$, because $\det(g)_v^w = 1$ if $g_v = 1$. We have $K^{(p)}/K_1 \cong K/(K_1 \times K_p)$, and the action of $K^{(p)}/K_1$ is independent of v. Using the action $\omega^{\otimes k} \mapsto g_{w,\varepsilon,v}\omega^{\otimes k}$, as

long as $\underline{\omega}^k_{K_1\times K_p/\mathcal{B}}$ is well-defined for $K \subset Z(\mathbb{A})S_0(\mathfrak{N})$, we can make a descent from $\underline{\omega}^k_{K_1\times K_p/\mathcal{W}}$ to $\underline{\omega}^k_{w,\varepsilon,K/\mathcal{W}}$, and we have for any $\mathcal{W}$-algebra $\mathcal{B}$,

$$H^0(Sh_K, \underline{\omega}^k_{w,\varepsilon,K/\mathcal{B}}) := H^0(K^{(p)}, H^0(Sh_{K_1\times K_p}, \underline{\omega}^k_{K_1\times K_p/\mathcal{B}})),$$

where the $H^0(K^{(p)}, ?)$ is with respect to the action $\omega^{\otimes k} \mapsto g_{w,\varepsilon,v}\omega^{\otimes k}$. We write this new sheaf as $\underline{\omega}^k_{w,\varepsilon,K}$ on Sh_K.

There are some other cases where the above definition yields $\underline{\omega}^k_{w,\varepsilon,K/\mathcal{B}}$ over Sh_K. If $v = \infty$ and $\mathcal{B} = \mathbb{C}$, the only requirement is to have smooth Sh_K. If $v = p$ and $\mathcal{B}$ is of characteristic p, $\underline{\omega}^k_{K_1\times K_p/\mathcal{B}}$ over $Ig_{K_1\times K_p}$ is well-defined, if $Ig/Ig_{K_1\times K_p}$ is étale and $K_p \supset U(\mathbb{Z}_p)$. Here we recall the unipotent radical U of the upper triangular Borel subgroup of G_1. Thus if $K \supset U(\mathbb{Z}_p)$ and Ig/Ig_K is étale ($\Leftrightarrow Ig_K$ is smooth), $\underline{\omega}^k_{w,\varepsilon,K/\mathcal{B}}$ is defined over Ig_K.

Abusing the language slightly, we call this Hecke character ε_+ *the central character* of $\underline{\omega}^k_{w,\varepsilon,K}$, because, as we show later in (4.55), modular forms f on $Sh_K(\mathbb{C})$ giving rise to global sections of $\underline{\omega}^k_{w,\varepsilon,K}$ satisfy $f|\langle z\rangle = \varepsilon_+(z)f$ for the normalized *diamond operator* $\langle z\rangle = |z|^{-2}_{\mathbb{A}}[KzK]$ with $z \in Z(\mathbb{A}^{(p\infty)})$ (the actual central character of f is $|\cdot|^{-2}_{\mathbb{A}}\varepsilon_+$). We define the v-type of the central character of the sheaf $\underline{\omega}^k_{w,\varepsilon,v,K}$ to be the character of $Z(\mathbb{Q})$ given by $\xi \mapsto \frac{\xi_{w,\varepsilon,v}\omega}{\varepsilon_+(\xi)\omega}$; so, we find the v-type to be $2w-k$. We now introduce the "new" double-digit weight $\kappa = (\kappa_1, \kappa_2) \in \mathbb{Z}[I]^2 = \mathrm{Hom}(T_G, \mathbb{G}_m)$ of $\underline{\omega}^k_{w,\varepsilon,v,K}$ by $\kappa_2 - \kappa_1 + I = k$ and $\kappa_1 + \kappa_2 = k - 2w + I$. Hereafter we denote this sheaf $\underline{\omega}^k_{w,v,K}$ by $\underline{\omega}^\kappa_{\varepsilon,K}$ and its cuspidal subsheaf by $\underline{\omega}_{\kappa,\varepsilon,K}$. As we have seen, these sheaves $\underline{\omega}^\kappa_\varepsilon$ and $\underline{\omega}_{\kappa,\varepsilon}$ satisfy the three axioms (s1–3). This change of our formulation of weights comes from a change of identification $T^2 \cong T_G$: (k,w) (resp. (κ_1,κ_2)) corresponds to $T^2 \ni (a,d) \mapsto d^{-1}\left(\begin{smallmatrix} a & 0 \\ 0 & a^{-1}\end{smallmatrix}\right)$ (resp. $T^2 \ni (a,d) \mapsto \left(\begin{smallmatrix} a & 0 \\ 0 & d\end{smallmatrix}\right)$), identifying T with $\mathrm{Res}_{O/\mathbb{Z}}\mathbb{G}_m$ (see (4.70) for the description of the automorphic factor of G giving rise to the sheaf $\underline{\omega}^\kappa_\varepsilon$).

Summarizing all these, we get

Theorem 4.24 *Let the notation be as above. We have*

(1) *If the base ring $\mathcal{B}$ is $\mathbb{C}$ and Sh_K is smooth, then $\underline{\omega}^\kappa_{\varepsilon,K}$ and its cuspidal subsheaf $\underline{\omega}_{\kappa,\varepsilon,K}$ are well-defined over $\mathbb{C}$;*

(2) *If K is maximal at p and $Sh^{(p)}_K$ is smooth, then $\underline{\omega}^\kappa_{\varepsilon,K}$ and its cuspidal subsheaf $\underline{\omega}_{\kappa,\varepsilon,K}$ are well-defined over $\mathcal{W}$;*

(3) *If K contains $U(\mathbb{Z}_p)$ and $Ig_{K/\mathcal{B}}$ is smooth for the base ring $\mathcal{B}$ of characteristic p or $\mathcal{B} = W$, then $\underline{\omega}^\kappa_{\varepsilon,K}$ and its cuspidal subsheaf $\underline{\omega}_{\kappa,\varepsilon,K}$ are well-defined over $Ig_{K/\mathcal{B}}$, where $Ig_{K/W}$ is the formal scheme $\varprojlim_m Ig_{K/W_m}$.*

The results of (1) and (2) in Theorem 4.24 give us a natural integral structure over $\mathcal{W}$ of the p-adic sheaf $\underline{\omega}^\kappa_{\varepsilon/W}$ and the Archimedean sheaf $\underline{\omega}^\kappa_{\varepsilon/\mathbb{C}}$ over $Sh^{(p)}_K$ (as long as they are well-defined). In other words, we have

$$\underline{\omega}^{\kappa}_{\varepsilon/\mathcal{W}} \otimes_{\mathcal{W}} W = \underline{\omega}^{\kappa}_{\varepsilon/W} \quad \text{and} \quad \underline{\omega}^{\kappa}_{\varepsilon/\mathcal{W}} \otimes_{\mathcal{W}} \mathbb{C} = \underline{\omega}^{\kappa}_{\varepsilon/\mathbb{C}} \tag{4.43}$$

at the same time. When ε restricted to $T(\widehat{\mathbb{Z}})$ is the identity character $\mathbf{1}$ and ε_+ is a power of the modulus character $|\cdot|_{\mathbb{A}}$, we drop ε from the notation, because then κ determines ε. As we observed already, when $k = mI$ ($m \in \mathbb{Z}$), we had a well-defined $\underline{\omega}^{mI}_{/\mathcal{W}}$, and we can easily verify that $\underline{\omega}^{mI}_{/\mathcal{W}} = \underline{\omega}^{\kappa}_{\varepsilon/\mathcal{W}}$ for $\kappa = (0, (m-1)I)$ and $\varepsilon_2 = \mathbf{1}$ and $\varepsilon_+(x) = |x|^{2-m}_{\mathbb{A}}$. In some of the earlier works of the author (e.g., [SGL] and [MFG]), $\Delta(\mathfrak{N})^{\iota}$ and $\varepsilon(x^{\iota})$ are used in place of $\Delta_0(\mathfrak{N})$, and also the roles of κ_1 and κ_2 are interchanged (particularly in the Hodge type of the motive associated to cusp forms of weight κ below). In other words, our formulation is cohomological (and hence Hecke operators act from the right, because cohomology functors are contravariant), while the formulation in [SGL] is homological; so, the Hecke operators act from the left. The two formulations are related by $[KgK'] \leftrightarrow [(KgK')^{\iota}]$, and the role of κ_1 and κ_2 is interchanged in the homological formulation (thus $(m, n) = (\kappa_2, \kappa_1)$ under the notation in [SGL]).

This double-digit weight κ has a geometric meaning that the motive M_f associated with a Hilbert modular Hecke eigenform $f \in H^0(Sh_K, \underline{\omega}_{\kappa,\varepsilon,K/\mathbb{C}})$ constructed by Blasius and Rogawski ([BlR] for $k > 2$, [H81] and [Bl] for $k = 2$) has Hodge weight $(\kappa_{2,\sigma}, \kappa_{1,\sigma})$ and $(\kappa_{1,\sigma}, \kappa_{2,\sigma})$ at the Archimedean place σ of F, and by writing $\widehat{\varepsilon}^{gal}_+$ for the ℓ-adic Galois character associated with the Hecke character ε_+, $\widehat{\varepsilon}^{gal}_+ \mathcal{N}_\ell$ with the ℓ-adic cyclotomic character $\mathcal{N}_\ell$ gives the determinant of the Galois representation of f (see [BlR], [Ta], and [MFG] 5.6.1 for a description of Galois representations and Section 4.2.7 in the text for $\widehat{\varepsilon}^{gal}_+$).

Our definition of ε^v_S in (ex3) looks technical; in particular, the extra factor $|z^{(v\infty)}|^2_{\mathbb{A}} z_v^{2I}$ in (4.41) appears to be strange, but it is placed in order to kill the factor $|\det(z)|_{\mathbb{A}}$ in (4.37) (and (4.44)) for the central element $z \in Z(\mathbb{A})$. This is why the central character is given by ε_+.

When $v = p$, by our definition of $w : T(\mathbb{Q}_p) \to W^{\times}$, g^* is well-defined for all g in the semi-group in (4.39). Still the multiplication by $|\det(g)|_{\mathbb{A}}$ in the definition of $[KgK']$ may not be well-defined in characteristic p (because $|p^{(\infty)}|_{\mathbb{A}}$ is not p-integral). We justify this later by means of the q-expansion principle.

We now relate the new automorphic definition of Hecke operators to the geometric one already given in Section 4.1.10. Let $K^1 = K \cap (G_1(\mathbb{Q}_v) \times G(\mathbb{A}^{(v)}))$. Decompose $K^1 g \cdot K'^1 = \bigsqcup_h K^1 h$ for $g \in \Delta_0((v) \cap \mathfrak{N})$, which yields $KgK' = \bigsqcup_h K \cdot h$. Then in the definition (4.37) of the operator $[KgK']$, the factor $\mathrm{Tr}_{K'/(K^g \cap K')} \circ [g]$ in down-to-earth terms is to average the image $h^*(f\omega^{\otimes k})$ over the coset representative h. Strictly speaking, this statement is only valid when the degree of the isogeny α_g is invertible over the base scheme S of the test object; otherwise, we need to stick to $\mathrm{Tr}_{K'/(K^g \cap K')}$ because $Sh_{K^g \cap K'}/Sh_{K'}$ (or $Ig_{K^g \cap K'}/Ig_{K'}$) could be inseparable. If $K \cdot h = K \cdot h'$, we have $(\eta \circ h^{-1}) \circ K = (\eta \circ h'^{-1}) \circ K$ as K-orbits. Thus we have a well-defined value

$f(A,\overline{\lambda},\overline{\eta\circ h^{-1}},\omega)$ for modular forms f of level K, because $(A,\overline{\lambda},\overline{\eta\circ h^{-1}},\omega)$ is a level K test object if so is $(A,\overline{\lambda},\overline{\eta},\omega)$. We suppose that $g\in\Delta_0(1)$. We recall our construction briefly: for a test object $(A,\lambda,\eta)_{/S}$, the subgroup $C=\eta(h^{-1}\widehat{L}/\widehat{L})$ for $\widehat{L}=\eta^{-1}(\mathcal{T}(A))$ is a subgroup of A defined over S. Let $Y_h=A/C$, and write $\alpha_h:A\to Y_h$ for the projection. We thus have (Y_h,λ_h,η_h) with $\eta_h=\alpha_h\circ\eta\circ h^{-1}$ and λ_h induced by α_h from λ. Write $x\in Sh(S)$ for the point represented by $(A,\overline{\lambda},\eta)$. Then xh^{-1} is represented by (Y_h,λ_h,η_h). Pick $\omega_x\in H^0(A,\Omega_{A/S})$. Then $\omega_{xh^{-1}}=h^{-1}\cdot\omega_x$ for the corresponding section over xh^{-1} is given by $\alpha_h^{-*}\omega_x$. For a global section $s\in H^0(Sh_K,\underline{\omega}^k)$ given by $s(x)=f(A,\overline{\lambda},\overline{\eta},\omega_x)\omega_x^{\otimes k}$ of $\underline{\omega}^k$ with a modular form f, we have

$$\begin{aligned}f|[KgK']_v(A,\overline{\lambda},\overline{\eta},\omega_x)\omega_x^{\otimes k}&=|\det(g)|_{\mathbb{A}}\sum_h h^*(f(A,\overline{\lambda},\overline{\eta}\circ h^{-1},\omega_{xh^{-1}})\omega_x^{\otimes k})\\&=\det(g_v)^w|\det(g)|_{\mathbb{A}}\sum_h\varepsilon_{\Delta}^v(h)f(Y_h,\overline{\lambda}_h,\overline{\eta}_h,\alpha_h^{-*}\omega_x)\omega_x^{\otimes k}.\quad(4.44)\end{aligned}$$

In the above formula, the summation symbol $\sum_h\varepsilon_{\Delta}^v(h)f(\underline{Y}_h)$ actually represents $\mathrm{Tr}_{K'/(K^g\cap K')}(\varepsilon_{\Delta}^v(g)f(\underline{Y}_g))$ if inseparability is involved (when α_g has noninvertible degree over the base of the test object) as we already remarked. The operator $[KgK']_v$ really depends on $\det(g_v)$ (because of the involvement of $(\det(g_v))^w\varepsilon_{\Delta}^v(h)$).

For an element z in the center with $z_\infty=1$, we define the diamond operator $\langle z\rangle_v=|z|_{\mathbb{A}}^{-2}[KzK']$ by

$$\begin{aligned}f|\langle z\rangle_v(A,\overline{\lambda},\overline{\eta},\omega_x)\omega_x^{\otimes k}&=|\det(z)|_{\mathbb{A}}\varepsilon_S^v(z)(f(A,\overline{\lambda},\overline{\eta}\circ z^{-1},\omega_{xz^{-1}})\omega_x^{\otimes k})\\&=\varepsilon_+^v(z)z_v^{-2I}f(Y_z,\overline{\lambda}_z,\overline{\eta}_z,(\alpha_z^*)^{-1}\omega_x)\omega_x^{\otimes k}.\quad(4.45)\end{aligned}$$

We have $\langle z\rangle_p=\langle z\rangle_\infty$ if $z_p=1$. It is customary to write $\langle\mathfrak{z}\rangle$ for $\langle z\rangle_v$ if $z\widehat{O}\cap F=\mathfrak{z}$ and $z_{p\mathfrak{N}}=1$ (in this case, the operator is independent of v).

If we take the shrinking semi-group $\Delta_0(\mathfrak{N})^\iota$ in place of our choice $\Delta_0(\mathfrak{N})$ here, the Hecke operator is induced by the right action of $g^\iota:Sh_K\to Sh_{K^{g^\iota}}$, and we need to replace the covering $(A,\overline{\lambda},\eta,\omega_{xh^{-1}})\xrightarrow{\alpha_g}(Y_h,\overline{\lambda}_h,\eta_h,(\alpha_h^*)^{-1}\omega_x)$ by another one $(Y'_{h^\iota},\overline{\lambda}'_{h^\iota},\eta'_{h^\iota},(\alpha'_{h^\iota})^*\omega_x)\xrightarrow{\alpha'_{h^\iota}}(A,\overline{\lambda},\eta,\omega_x)$ for the operator $[(KgK')^\iota]$ in (4.44) (or simply to make the following variable change, $g\mapsto g^{-\iota}$, as was done in [H88], page 307). We have made our choice of the cohomological action, because it is directly related to the classical Hecke operator $T(\mathfrak{q})$ already defined in Section 4.1.10.

We assume that $S_1^1(\mathfrak{N})\subset K\subset S_0(\mathfrak{N})$. If we consider a test object $(A,\overline{\lambda},\overline{\eta})$ sitting over a point of $S^\circ_{K,\infty}=Sh_K[\frac{1}{E}]$, we have the connected component $C_0(\mathfrak{p})$ of $A[p]$, and all finite subgroups C of A with $C\cap C_0(\mathfrak{p})=C\times_A C_0(\mathfrak{p})=\{\mathbf{0}\}$ (scheme-theoretically) are étale. Thus if the test object is p-ordinary and sitting over $T_{K,m,\alpha}$ for $\alpha>0$, we may normalize L to be given by $O^*\oplus\mathfrak{c}^{-1}\subset V$ with $\eta(\widehat{L})=\mathcal{T}(A)$, and we may assume (by

the irreducibility of $T_{K,m,\alpha}$ proven in Theorem 4.21) that $\eta(\mathfrak{p}^*/O^*) = C_0(\mathfrak{p})$ at the characteristic 0 fiber of A. For any prime factor $\mathfrak{p}$, taking the uniformizer $\varpi_\mathfrak{p} = p$ of $O_\mathfrak{p}$, we find $K_0(\mathfrak{p})^1 \left(\begin{smallmatrix} 1 & 0 \\ 0 & \varpi_\mathfrak{p} \end{smallmatrix}\right) K_0(\mathfrak{p})^1$ for $K_0(\mathfrak{p}^r) = K \cap \widehat{\Gamma}_0(\mathfrak{p}^r)$ is decomposed into $\bigsqcup_{u \bmod \mathfrak{p}} K_0(\mathfrak{p})^1 h_u$ for $h_u = \left(\begin{smallmatrix} 1 & u \\ 0 & \varpi_\mathfrak{p} \end{smallmatrix}\right)$. Thus all étale cyclic subgroups of order $\mathfrak{p}$ in A are given by $\eta(h_u^{-1}\widehat{L}/\widehat{L})$, and the definition of the Hecke operator $U(\mathfrak{p})$ in (4.16) is identical to the operator $[K_0(\mathfrak{p}) \left(\begin{smallmatrix} 1 & 0 \\ 0 & \varpi_\mathfrak{p} \end{smallmatrix}\right) K_0(\mathfrak{p})]_\infty = [K_0(\mathfrak{p}) \left(\begin{smallmatrix} 1 & 0 \\ 0 & \varpi_\mathfrak{p} \end{smallmatrix}\right) K_0(\mathfrak{p})]_p$ given by (4.44) (because $\varpi_\mathfrak{p}^{-\kappa_1} = 1$ by (ex0)). We write $T_v(1,y)$ for $[K \left(\begin{smallmatrix} 1 & 0 \\ 0 & y \end{smallmatrix}\right) K]_v$, which really depends on $y \in (\mathbb{A}^{(\infty)})^\times$ (as we already remarked). We have $T_p(1,y) = y_p^{-\kappa_1} T_\infty(1,y)$. When $S_1^1((p) \cap \mathfrak{N}) \subset K \subset S_0((p) \cap \mathfrak{N})$ and $y^{(p\mathfrak{N})} = 1$, we write $U_v(y)$ for $T_v(1,y)$. If $y_p = 1$, these operators do not depend on v; so, we drop v from our symbols if the choice v is clear from the context. Then we have

$$U_v(\varpi_\mathfrak{p})^n = U_v(\varpi_\mathfrak{p}^n) \quad \text{if } n > 0. \tag{4.46}$$

Note here that $\varpi_\mathfrak{p}'^{-\kappa_1} \in W^\times$ by the definition of $\varpi_\mathfrak{p}'^{-\kappa_1}$ in (ex0) for any other generator $\varpi_\mathfrak{p}'$ of $\mathfrak{p}O_\mathfrak{p}$, and hence

$$e = \prod_{\mathfrak{p}|p} e_\mathfrak{p} \quad \text{with} \quad e_\mathfrak{p} = \lim_{n\to\infty} U_p(\varpi_\mathfrak{p}')^{n!} = \lim_{n\to\infty} U_p(\varpi_\mathfrak{p})^{n!} = \lim_{n\to\infty} U(\mathfrak{p})^{n!}.$$

When a prime $\mathfrak{q}|\mathfrak{N}$, we have a specific subgroup $C_0(\mathfrak{q}) = \eta((\mathfrak{c}\mathfrak{q})^*/\mathfrak{c}^*)$, and recalling the uniformizer $\varpi_\mathfrak{q}$ of $O_\mathfrak{q}$, we find

$$U_v(\varpi_\mathfrak{q}^n) = [K \left(\begin{smallmatrix} 1 & 0 \\ 0 & \varpi_\mathfrak{q} \end{smallmatrix}\right) K]_v^n = [K \left(\begin{smallmatrix} 1 & 0 \\ 0 & \varpi_\mathfrak{q}^n \end{smallmatrix}\right) K]_v \quad \text{for } v = p \text{ and } \infty. \tag{4.47}$$

If $\mathfrak{q}$ is outside $p\mathfrak{N}$, we have

$$T(1, \varpi_\mathfrak{q}^n) = [K \left(\begin{smallmatrix} 1 & 0 \\ 0 & \varpi_\mathfrak{q}^n \end{smallmatrix}\right) K]_v \quad \text{for } v = p \text{ and } \infty. \tag{4.48}$$

The operator $T_v(1,y)$ with $y_p = 1$ does not involve $y_p^{-\kappa_1}$ and $|y_p|_\mathbb{A} = |\det(\mathrm{diag}[1,y])|_\mathbb{A}$; so, it is integral over $\mathcal{W}$. We record what we have proven.

Corollary 4.25 *Let the notation and the assumption be as in* Theorem 4.24 (2). *Suppose that* $\varepsilon^- = \varepsilon_2^{-1}\varepsilon_1$ *factors through* $(O/\mathfrak{N})^\times$. *Then we have* $\mathcal{W}$*-integral line bundles* $\underline{\omega}_{\kappa,\varepsilon} \subset \underline{\omega}_\varepsilon^\kappa$ *satisfying the three axioms* (s1–3) *with an action of* $\Delta_0((p) \cap \mathfrak{N})$ *in* (s2) *specified in* (4.42) *for the coefficient rings* W *and* $\mathbb{C}$. *These line bundles descend to* $Sh^{(p)}_{K/\mathcal{W}}$ *for* $K \subset \Delta_0(\mathfrak{N})$ *maximal at* p *as long as* $Sh^{(p)}/Sh^{(p)}_K$ *is étale. The Hecke operators* $T_p(1,y) = T_\infty(1,y)$ *with* $y_p = 1$ *are* $\mathcal{W}$*-integral.*

In Theorem 4.28, we describe when W-integrality for $T_p(1,y)$ and $U_p(y)$ with nontrivial y_p holds.

4.2.7 Sheaves over the Shimura Variety of $PGL(2)$

We have constructed $\underline{\omega}^{\kappa}_{\varepsilon,K}$ for a closed subgroup K of $\Delta = \Delta_0((p) \cap \mathfrak{N})$. If $K^{(p)}$ contains the center $Z(\mathbb{A}^{(p\infty)})$, the center acts on the sheaf $\underline{\omega}^{\kappa}_{\varepsilon,K}$ by the character ε^p_S, and we may regard $\underline{\omega}^{\kappa}_{\varepsilon}$ as a line bundle defined over the Shimura variety $Sh(PG, X)$ of the group $PG = \mathrm{Res}_{O/\mathbb{Z}}PGL(2)$, because the central action is specified. In this subsection, we construct a slightly larger family of sheaves on $Sh(PG, X)$ containing all $\underline{\omega}^{\kappa}_{\varepsilon}$ so far constructed.

We interpret the sheaf $\underline{\omega}^{\kappa}_{\varepsilon}$ as a sheaf over $Sh(PG, X)$. For $a \in Z(\mathbb{A}^{(\infty)})$, writing $\mathfrak{a}$ for $F \cap a\widehat{O}$, we find $\mathcal{T}(A \otimes \mathfrak{a}) = \mathcal{T}(A) \otimes_{\mathbb{Z}} \mathfrak{a}$ and $\eta^{-1}(\mathcal{T}(A \otimes \mathfrak{a})) = a\eta^{-1}(\mathcal{T}(A))$ for a test object (A, λ, η). Tensoring A with $\mathfrak{a} \subset F \supset O$, two maps $A \otimes \mathfrak{a} \to A \otimes F \leftarrow A$ induce an isogeny $\alpha \in \mathrm{Hom}^{\mathbb{Q}}_F(A \otimes \mathfrak{a}, A)$. We get $(A, \overline{\lambda}, \eta \circ a) \cong (A \otimes \mathfrak{a}, \alpha^*\overline{\lambda}, \alpha^{-1} \circ \eta)$. This central action brings the $\mathfrak{c}$-polarization to $\mathfrak{c}\mathfrak{a}^2$-polarization $\alpha^*\overline{\lambda}$ as seen in Section 4.1.9.

Since the Shimura variety $Sh(PG, X)$ is given by $Sh(G, X)/Z(\mathbb{A}^{(\infty)})$, we identify $(A, \overline{\lambda}, \eta \circ a, \omega) = (A \otimes \mathfrak{a}, \alpha^*\overline{\lambda}, \alpha^{-1} \circ \eta, \alpha^*\omega)$ with $(A, \overline{\lambda}, \eta)$. If $\underline{\omega}^{\kappa}_{\varepsilon}$ descends to $Sh(PG, X)$, the center $Z(\mathbb{A}^{(\infty)})$ has to act trivially on $\underline{\omega}^{\kappa}_{\varepsilon}$; so, we find that $k - 2w = 0$ (hence $k \in 2\mathbb{Z}[I]$) and that ε_+ is trivial. This means $\kappa = (I - w, w)$. We write the sheaf $\underline{\omega}^{(I-w,w)}_{\varepsilon}$ as $\underline{\omega}^{w}_{\varepsilon}$ for $w \in \mathbb{Z}[I]$. We regard $\underline{\omega}^{w}_{\varepsilon}$ as defined over $Sh(PG, X) = Sh_{Z(\mathbb{A}^{(\infty)})}(G, X)$.

Fix the Neben character ε with trivial ε_+, and consider $\underline{\omega}^{w}_{\varepsilon}$ as specified above. We can generalize the above construction slightly allowing nontrivial central characters χ. As before, we write $\mathcal{B}$ for the base ring (either W or $\mathbb{C}$). The classification functor of $\underline{A} = (A, \overline{\lambda}, \overline{\eta}^{(p)}, \omega)$,

$$\mathcal{P}^{(p)}_1(S) = \left[\underline{A}_{/S} \middle| (A, \overline{\lambda}, \overline{\eta}^{(p)}) \in \mathcal{E}^{(p)}_1(S),\ H^0(A, \Omega_{A/S}) = (O \otimes \mathcal{O}_S)\omega\right],$$

is represented by a T-torsor $\mathcal{M}^{(p)}_{\{1\}}$ over $Sh(G, X)^{(p)}$. Composing

$$\mathcal{M}^{(p)}_{\{1\}} \to Sh^{(p)}(G, X) \to Sh^{(p)}_{Z(\mathbb{Z}_p \times \mathbb{A}^{(p\infty)})}(G, X),$$

we may regard $\mathcal{M}^{(p)}_{\{1\}}$ as a $\mathcal{G}$-torsor over the variety $Sh^{(p)}_{Z(\mathbb{A}^{(\infty)})}(G, X)$ for $\mathcal{G} = (Z(\mathbb{Z}_p \times \mathbb{A}^{(p\infty)})/\overline{Z(\mathbb{Z}_{(p)})}) \times T$. Here $Z(\mathbb{Z}_p \times \mathbb{A}^{(p\infty)})/\overline{Z(\mathbb{Z}_{(p)})}$ is a constant group. Take the vector bundle $\underline{\Omega}$ associated with this torsor over $Sh^{(p)}_{Z(\mathbb{A}^{(\infty)})}(G, X)$. For each character $(\chi, 2w) : \mathcal{G} \to \mathcal{B}^{\times}$ sending (z, x) to $\chi(z)x^{2w}$ ($x \in T$ and $z \in (Z(\mathbb{Z}_p \times \mathbb{A}^{(p\infty)})/\overline{Z(\mathbb{Z}_{(p)})})$), we can cut out a line bundle $\underline{\omega}^{w}_{\varepsilon}[\chi]_{/\mathcal{B}}$ given by the sheaf of eigensections under character $\mathcal{G} \ni (z, x) \mapsto \chi^{-1}(z)x^{-2w}$. In other words, we let the center $Z(\mathbb{A}^{(p\infty)})$ act on $\underline{\omega}^{w}_{\varepsilon/Sh^{(p)}(G,X)}$ by χ and descend the sheaf $\underline{\omega}^{w}_{\varepsilon/Sh^{(p)}(G,X)}$ to $\underline{\omega}^{w}_{\varepsilon}[\chi]_{/\mathcal{B}}$ over $Sh^{(p)}(PG, X)$ by this new action.

Since $Z(\mathbb{A}^{(\infty)}) \supset \overline{Z(\mathbb{Q})} \cap Z(\mathbb{Z}_p \times \mathbb{A}^{(p\infty)}) = \overline{Z(\mathbb{Z}_{(p)})} \subset Z(\mathbb{Z}_p \times \mathbb{A}^{(p\infty)})$, we have

$$\frac{Z(\mathbb{Z}_p \times \mathbb{A}^{(p\infty)})}{\overline{Z(\mathbb{Z}_{(p)})}} \cong \frac{Z(\mathbb{Z}_p \times \mathbb{A}^{(p\infty)})\overline{Z(\mathbb{Q})}}{\overline{Z(\mathbb{Q})}} = \frac{Z(\mathbb{A}^{(\infty)})}{\overline{Z(\mathbb{Q})}} = \frac{Z(\mathbb{A})}{\overline{Z(\mathbb{Q})Z(\mathbb{R})}}.$$

Thus the characters of $\frac{Z(\mathbb{Z}_p \times \mathbb{A}^{(p\infty)})}{\overline{Z(\mathbb{Z}_{(p)})}}$ are in bijection with continuous even characters of the Galois group $\mathrm{Gal}(F^{ab}/F) \cong \frac{Z(\mathbb{A})}{\overline{Z(\mathbb{Q})Z(\mathbb{R})^+}}$ by the Artin reciprocity law of class field theory. A continuous Hecke character $\chi : Z(\mathbb{A})/Z(\mathbb{Q}) = F_{\mathbb{A}}^\times/F^\times \to \mathbb{C}^\times$ is called *arithmetic* (in [MFG] 1.2.2) if we have $w' \in X(Z) = \mathbb{Z}[I]$ such that $\chi(x_\infty) = x_\infty^{-w'}$ whenever x_∞ is totally positive. Suppose that χ is arithmetic and that the base algebra $\mathcal{W}$ is so large that $w' \in X(Z)$ is defined over $\mathcal{W}$ (i.e., $x^{w'} \in \mathcal{W}^\times$ if $x \in Z(\mathcal{W})$). We have $\chi(x^{(\infty)}) \in \overline{\mathbb{Q}}^\times \subset \overline{\mathbb{Q}}_p^\times$ for $x^{(\infty)} \in Z(\mathbb{A}^{(\infty)})$. Then we define a p-adically continuous character $\widehat{\chi} : Z(\mathbb{A}^{(\infty)})/Z(\mathbb{Q})^+ = {F_{\mathbb{A}^{(\infty)}}}^\times/F_+^\times \to \overline{\mathbb{Q}}_p^\times$ by $\widehat{\chi}(x) = i_p(\chi(x^{(\infty)}))x_p^{-w'}$ (see [MFG] Theorem 1.1). By the Artin reciprocity map, we can regard $\widehat{\chi}$ as a Galois character $\widehat{\chi}^{gal}$. Thus if $\widehat{\chi}$ is even, we have well-defined $\underline{\omega}_\varepsilon^w[\chi] := \underline{\omega}_\varepsilon^w[\widehat{\chi}]$ as an invertible sheaf over $Sh(PG, X)$. The central character of $\underline{\omega}_\varepsilon^w[\chi]$ is given by $z \mapsto \chi(z^{(p)})^{-1}z_p^{w'}$. Since the weight of $\underline{\omega}_\varepsilon^\kappa$ is determined by its character ε_+ and κ, we may define the weight of $\underline{\omega}_\varepsilon^w[\chi]$ to be $\kappa = (\frac{w'}{2} + I - w, \frac{w'}{2} + w)$. Choosing $\varepsilon_\chi : T_G(\widehat{\mathbb{Z}}) \to \mathcal{W}^\times$ with $\varepsilon_{\chi+} = \chi$, we descend $\underline{\omega}_\varepsilon^w[\chi]$ to $\underline{\omega}_{\varepsilon\varepsilon_\chi,K}^\kappa$ over $Sh_{\overline{K}}^{(p)}$ if $w' \in 2\mathbb{Z}[I]$ for K with $S_1^1(\mathfrak{N}) \subset K \subset S_0(\mathfrak{N})$, where $\overline{K} = Z(\mathbb{A}^{(\infty)})K$ (so, $\overline{K}/Z(\mathbb{A}^{(\infty)})$ is an open compact subgroup of $PG(\mathbb{A}^{(\infty)})$ maximal at p).

When $\chi = \psi^2$, we can interpret the central twist $\underline{\omega}_\varepsilon^w[\chi]$ of $\underline{\omega}_\varepsilon^w$ as an external twist $\underline{\omega}_\varepsilon^w \otimes \psi$. We now introduce a twist of the sheaf $\underline{\omega}_\varepsilon^\kappa$ over $Sh(PG, X)$ by the square root ψ of χ. The pullback of $\underline{\omega}_\varepsilon^\kappa$ to $Sh(G, X)$ has the natural action $g_{w,\varepsilon}$ of $g \in \Delta$ in (s2) as we already specified. Using determinant character $\det : G \to T$, we can twist the action $g_{w,\varepsilon}$ on $\underline{\omega}_\varepsilon^\kappa$ by $\omega^{\otimes k} \mapsto g_{w,\varepsilon,\chi}\omega^{\otimes k} := \psi^{-1}(\det(g)^{(\infty)})g_{w,\varepsilon}\omega^{\otimes k}$. We denote this sheaf as $\underline{\omega}_\varepsilon^\kappa \otimes \psi$. This sheaf $\underline{\omega}_\varepsilon^\kappa \otimes \psi$ has the central character $\chi\varepsilon_+$, and locally, $u = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in S_0(\mathfrak{N})$ for the conductor $\mathfrak{N}$ of ε^- acts on the sheaf by $\varepsilon^-(a_\mathfrak{N})\varepsilon_2\psi(ad - bc)$. Defining $\varepsilon \otimes \psi$ by $(\varepsilon_1\psi^{(\infty)}, \varepsilon_2\psi^{(\infty)})$, we find $(\varepsilon \otimes \psi)_+ = \varepsilon_+\chi$ and

$$\underline{\omega}_\varepsilon^\kappa \otimes \psi \cong \underline{\omega}_{\varepsilon\otimes\psi}^{\kappa+(w'/2,w'/2)};\ \text{particularly, } \underline{\omega}_\varepsilon^\kappa \otimes \psi = \underline{\omega}_\varepsilon^w[\chi] \text{ if } \varepsilon_+ \text{ is trivial.} \tag{4.49}$$

Here is a criterion for the existence of $\psi = \sqrt{\chi}$.

Lemma 4.26 *If a finite-order continuous character $\chi : F_{\mathbb{A}}^\times/F^\times \to \overline{\mathbb{Q}}^\times$ factors through $F_{\mathbb{A}}^\times/F^\times F_\infty^\times = F_{\mathbb{A}^{(\infty)}}^\times/F^\times$, then there exists a continuous character $\sqrt{\chi} : F_{\mathbb{A}}^\times/F^\times \to \mathbb{C}^\times$ with $\sqrt{\chi}^2 = \chi$.*

Proof. Suppose that $\xi^2 = \eta \in F^\times$ for $\xi \in F_{\mathbb{A}^{(\infty)}}$. Then, all primes of F split in $F[\sqrt{\eta}]$; so, $F = F[\sqrt{\eta}]$ by class field theory. Thus $F_{\mathbb{A}^{(\infty)}}^\times/F^\times F_\infty^\times$ is 2-torsion-free; so, we can find a continuous character $\sqrt{\chi}$ as above. □

A character of $\mathrm{Gal}(F^{ab}/F)$ is called *totally even* if it is trivial on any complex conjugation (i.e., $i^{-1} \circ c \circ i$ for any embedding $i : F^{ab} \hookrightarrow \mathbb{C}$). If a Galois character $\widehat{\chi} : \mathrm{Gal}(F^{ab}/F) \to \mathcal{B}^\times$ is totally even, the corresponding idele character factors through $F^\times_{\mathbb{A}^{(\infty)}}/F^\times$, and we find a character $\psi : \mathrm{Gal}(F^{ab}/F) \to \mathcal{B}^\times$ with $\widehat{\chi} = \psi^2$.

4.2.8 Hecke Algebra of Finite Level

Let R be a $\mathcal{W}$-algebra. Fix a Neben character $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_+)$ with values in $R^\times$ as in Theorem 4.24 and a dominant regular weight $\kappa = (\kappa_1, \kappa_2)$ in $X(T_G)$. The weight is dominant regular if $\kappa_2 > \kappa_1$ and is just dominant if $\kappa_2 \geq \kappa_1$. We define, for $\mathfrak{N}$ prime to p,

$$G_\kappa(\mathfrak{N}, \varepsilon; R) = H^0(Sh^{(p)}_{\overline{S}_0(\mathfrak{N})/R}, \underline{\omega}^\kappa_{\varepsilon/R}) \text{ if } \kappa = (\mathbb{Z} \cdot I)^2,$$
$$\text{and } S_\kappa(\mathfrak{N}, \varepsilon; R) = H^0(Sh^{(p)}_{\overline{S}_0(\mathfrak{N})/R}, \underline{\omega}_{\kappa,\varepsilon/R}), \quad (4.50)$$

where $\overline{K} = Z(\mathbb{A}^{(\infty)})K$ (so, $\overline{S}_0(\mathfrak{N}) = Z(\mathbb{A}^{(\infty)})S_0(\mathfrak{N})$). The group $\overline{K}/Z(\mathbb{A}^{(\infty)})$ is an open compact subgroup of $PG(\mathbb{A}^{(\infty)})$. We then define the Hecke algebra $h_\kappa(\mathfrak{N}, \varepsilon; R)$ with central character ε_+ and Neben character ε to be the R-subalgebra of $\mathrm{End}_R(S_\kappa(\mathfrak{N}, \varepsilon; R))$ generated by $T_\infty(1, \varpi_\mathfrak{q})$ for all primes $\mathfrak{q}$ outside $\mathfrak{N}$ and $U_\infty(y)$ (with all $y \in F^\times_\mathfrak{N} \cap O_\mathfrak{N}$) for all primes $\mathfrak{q}|\mathfrak{N}$.

For more general K maximal at p, we define

$$G_\kappa(K, \varepsilon; R) = H^0(Sh_{\overline{K}/R}, \underline{\omega}^\kappa_{\varepsilon/R}) \text{ if } \kappa \in (\mathbb{Z} \cdot I)^2,$$
$$\text{and } S_\kappa(K, \varepsilon; R) = H^0(Sh_{\overline{K}/R}, \underline{\omega}_{\kappa,\varepsilon/R}). \quad (4.51)$$

Since $\overline{K}$ contains all the information of the center, the operator $\langle z \rangle$ ($z \in Z(\mathbb{A}^{(p\infty)})$) acts on $G_\kappa(K, \varepsilon; R)$ through the scalar multiplication by $\varepsilon_+(z)$ (see (4.55)).

For K with $S^1_1(\mathfrak{N}) \subset K \subset S_0(\mathfrak{N})$, we define the Hecke algebra $h_\kappa(K, \varepsilon; R)$ with coefficients in R by the R-subalgebra of $\mathrm{End}_R(S_\kappa(K, \varepsilon; R))$ in the same manner as above.

In the following subsection, we show that the operators $T_\infty(1, \varpi_\mathfrak{q})$ and $U_\infty(\varpi_\mathfrak{q})$ are $\mathcal{W}$-integral for almost all cases, at least if $\kappa_1 + \kappa_2 \geq I$ and κ is dominant; see (h1–6) for more details.

More generally, for a subgroup K with $(S_1(p^r) \cap S^1_1(\mathfrak{N})) \subset K \subset S_0(p^r\mathfrak{N})$ with $r > 0$, we define for a profinite W-algebra R,

$$G_\kappa(K, \varepsilon; R) = H^0(Ig_{\overline{K}/R}, \underline{\omega}^\kappa_{\varepsilon/R}) \text{ if } \kappa = (\mathbb{Z} \cdot I)^2$$
$$\text{and } S_\kappa(K, \varepsilon; R) = H^0(Ig_{\overline{K}/R}, \underline{\omega}_{\kappa,\varepsilon/R}). \quad (4.52)$$

Here $Ig_{\overline{K}/R}$ is regarded as the quotient of the Igusa tower $Ig_{K/R}$ by the action of $Z(\mathbb{Z}_p \times \mathbb{A}^{(p)})$. Since Ig_K is an open formal scheme over W, these spaces may not be of finite type over R. However, its ordinary part is an R-module of finite type (Theorem 4.11).

Proposition 4.27 *Suppose that κ is a dominant weight of T_G. If K is maximal at p, $S_\kappa(K,\varepsilon;R)$ and $G_\kappa(K,\varepsilon;R)$ are R-modules of finite type.*

Proof. Since κ is dominant, $\mathcal{L} = \underline{\omega}^\kappa_{\varepsilon/R}$ for $\kappa \in (\mathbb{Z}\cdot I)^2$ or $\mathcal{L} = \underline{\omega}_{\kappa,\varepsilon/R}$ for general κ extends canonically to a line bundle over the smooth projective toroidal compactification Sh^{sm}_K of $Sh^{(p)}_K$. By the Koecher principle (see (4.10)), we have $H^0(Sh^{(p)}_{\overline{K}/R},\mathcal{L}) = H^0(Sh^{sm}_{\overline{K}/R},\mathcal{L})$ for $\mathcal{L}$ as above if $\kappa \in (\mathbb{Z}\cdot I)^2$. Taking $\kappa' = (0,(m-1)I)$ for a sufficiently large m, $\underline{\omega}^{\kappa'-\kappa}_{\varepsilon'/R}$ for $\varepsilon' = (|\cdot|^{-m}_{\mathbb{A}}, \mathbf{1}, |\cdot|^{m-2}_{\mathbb{A}})$ has a global section s over Sh^{sm}_K such that the multiplication by s gives an injection of $S_\kappa(K,\varepsilon;R)$ into $S_{\kappa'}(K,\varepsilon\varepsilon';R)$, because $\underline{\omega}^{mI}_{\varepsilon'}$ is very ample for m large. This shows that these spaces are R-modules of finite type (e.g., [ALG] III.8.8). □

We define $h_\kappa(K,\varepsilon;R)$ and $h_\kappa(p^r\mathfrak{N},\varepsilon;R)$ by the R-profinite subalgebra of $\mathrm{End}_R(S_\kappa(K,\varepsilon;R))$ topologically generated by $T_\infty(1,\varpi_\mathfrak{q})$ for all primes $\mathfrak{q}$ outside $\mathfrak{N}$ and $U_\infty(\varpi_\mathfrak{q})$ for all primes $\mathfrak{q}|\mathfrak{N}$.

4.2.9 Effect on q-Expansion

If p is unramified in $F/\mathbb{Q}$, by Theorem 4.21, $\mathfrak{M}(\mathfrak{c},\Gamma)_{/\mathcal{W}}$ and the Igusa tower T_{∞/W_m} over it are geometrically irreducible; so, the q-expansion $f(q) = f(\mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q))$ determines the section f of $\underline{\omega}^\kappa_{\varepsilon/R}$ for any $\mathcal{W}$-algebra R (see Corollary 4.23). This fact we refer to the q-expansion principle over $\mathcal{W}$. Here, for a given open compact subgroup $K \supset S^1_1(\mathfrak{N})$ (giving rise to the level structure of type Γ), $\mathcal{W}$ is assumed to be a sufficiently large discrete valuation ring inside $\overline{\mathbb{Q}}$ giving the p-adic place induced by i_p over which the generically irreducible component of $Sh^{(p)}_{K/\mathcal{W}}$ is fiber-by-fiber geometrically irreducible. We recall $W = \varprojlim_n \mathcal{W}/p^n\mathcal{W}$.

We compute the q-expansion of $f|t$ for Hecke operators t in order to show that the Hecke operators preserve $\mathcal{W}$-integrality. When we do this computation in the classical setting of $\underline{\omega}^k$ over $\mathfrak{M}(\mathfrak{c},\Gamma)$, we assume $f \in S_k(\mathfrak{c},\Gamma;R)$. When we deal with adelic Hecke operators $T_v(1,y)$ and $U_v(y)$, we assume that $f \in S_\kappa(K,\varepsilon;R)$ for $K \supset S^1_1(\mathfrak{N})$; so, in particular, f has the central character ε_+. We prove

Theorem 4.28 *Let the notation be as above, and assume that p is unramified in $F/\mathbb{Q}$. Let $\mathfrak{N}$ be an O-ideal prime to p and R be a W-algebra.*

First let K be a subgroup with $S^1_1(p^r\mathfrak{N}) \subset K \subset S_0(p^r\mathfrak{N})$ for $r > 0$. Then the operators $[KgK]_p$ for $g \in \Delta_0(\mathfrak{N})$ with $g_p = 1$, $\langle z\rangle_p$ for $z \in Z(\mathbb{A}^{(p\infty)})$, $T_p(1,y)$ and $U_p(y)$ ($y \in F^\times_{\mathbb{A}} \cap \widehat{O}$) are well-defined on $G_\kappa(K,\varepsilon;R)$ and $S_\kappa(K,\varepsilon;R)$ if one of the following conditions is satisfied.

(h1) *$R = \mathcal{B}/p^m\mathcal{B}$ for a W-algebra $\mathcal{B}$, and the Hecke operators are well-defined for $G_\kappa(K,\varepsilon;\mathcal{B})$ (if $\kappa \in (\mathbb{Z}\cdot I)^2$), and $S_\kappa(K,\varepsilon;\mathcal{B})$ (for all κ);*
(h2) *$R = W$ or $W_m = \mathcal{W}/p^m\mathcal{W}$ and $(S_1(p^r)\cap S^1_1(\mathfrak{N})) \subset K \subset S_0(p^r\mathfrak{N})$;*

(h3) *R is p-torsion-free.*

Suppose that $S_1^1(\mathfrak{N}) \subset K \subset S_0(\mathfrak{N})$. Then the operators $\langle z\rangle_\infty$ for $z \in Z(\mathbb{A}^{(\infty)})$, $[KgK]_\infty$ for $g \in \Delta_0(\mathfrak{N})$ with $g_p = 1$, and $T_\infty(1,\mathfrak{p})$ are R-integral under one of the following conditions.

(h4) *We have $R = \mathcal{B}/p^m\mathcal{B}$ for a W-algebra $\mathcal{B}$, the Hecke operators are well-defined for $G_\kappa(K,\varepsilon;\mathcal{B})$ and $S_\kappa(K,\varepsilon;\mathcal{B})$, and $\underline{\omega}_\varepsilon^\kappa$ is very ample on the minimal compactification $Sh^*_{K/\mathcal{B}}$ of $Sh_{K/\mathcal{B}}$;*

(h5) *$R = W$ or $W_m = \mathcal{W}/p^m\mathcal{W}$, $\kappa_1 + \kappa_2 \geq 0$, and $S_1^1(\mathfrak{N}) \subset K \subset S_0(\mathfrak{N})$;*

(h6) *R is a $\mathbb{Q}$-algebra.*

Under (h5)*, if we further assume $\kappa_1 + \kappa_2 \geq I$, we have $T_\infty(1,\varpi_\mathfrak{p}) \equiv U_p(\varpi_\mathfrak{p})$ mod p in $h_\kappa(K \cap S_0((p)),\varepsilon;W)$, where $U_p(\varpi_\mathfrak{p})$ is defined on the bigger space $G_\kappa(K \cap S_0(p),\varepsilon;W)$ than $G_\kappa(K,\varepsilon;W)$ on which $T_\infty(1,\varpi_\mathfrak{p})$ is defined.*

Proof. We first treat the case where (h4) is valid. Since $\underline{\omega}_\varepsilon^\kappa$ is very ample, it is generated by global sections, and $G_\kappa(K,\varepsilon;\mathcal{B}) \otimes_\mathcal{B} R = G_\kappa(K,\varepsilon;R)$. Since $G_\kappa(K,\varepsilon;\mathcal{B})$ is stable under the Hecke operators, $G_\kappa(K,\varepsilon;R)$ is also kept by the Hecke operators. Once we know that G_κ is stable under the Hecke operators, S_κ is also stable, because the cusps are permuted by the action of $\underset{\sim}{g} \in \Delta$.

To see the assertion under (h1), we consider the normalization Ig_K of the minimal compactification M_K^* of $Sh^{(p)}_{K^{(p)}\times G(\mathbb{Z}_p)}$ in Ig_K. Let S_K^* be the formal completion of $M_K^*\left[\frac{1}{E}\right]$ along the modulo p special fiber for a lift E of the Hasse invariant. Then we put $Ig_K^* = \widetilde{Ig}_K \times_{M_K^*} S_K^*$. Then S_K^* and Ig_K^* are affine schemes, and Ig_K^* is étale over S_K^*. This is because the Igusa tower Ig_K^{sm} over the smooth toroidal compactification $Sh^{sm}_{K^{(p)}\times G(\mathbb{Z}_p)}$ is étale and the infinitesimal structure of the projection $Ig_K^{sm} \to Ig_K^*$ around each cusp of Ig_K^* is isomorphic (infinitesimally) to that of the projection $Sh^{sm}_{K^{(p)}\times G(\mathbb{Z}_p)} \to M_K^*$. Note that $\underline{\omega}_\varepsilon^\kappa$ for $\kappa \in (\mathbb{Z}\cdot I)^2$ and $\underline{\omega}_{\varepsilon,\kappa}$ extend to Ig_K^* and $H^0(Ig_K^*,\mathcal{L}) = H^0(Ig_K^{sm},\mathcal{L})$ by the Koecher principle for $\mathcal{L} = \underline{\omega}_\varepsilon^\kappa$ and $\underline{\omega}_{\varepsilon,\kappa}$. Since Ig_K^* is affine, we find $\mathcal{F}_\kappa(K,\varepsilon;\mathcal{B}) \otimes_\mathcal{B} R = \mathcal{F}_\kappa(K,\varepsilon;R)$ for $\mathcal{F} = G$ and S. Then we conclude the assertion in the same way as in the case of (h4).

The operators may not be well-defined, because of the factor $|\det(g)|_\mathbb{A}$ in (4.37), (4.44), and (4.45) (see also (4.16) and (4.17)). If R is a $\mathbb{Q}$-algebra, this does not cause any problem; so, the assertion under (h6) follows.

To show the assertion under (h2), (h3), or (h5), we apply the q-expansion principle as described at the beginning of this subsection and Corollary 4.23 to verify that $f|[KgK]$ and $f|\langle z\rangle$ are R-integral as long as f is R-integral. We first compute the q-expansion of $f|\langle\mathfrak{z}\rangle$ for $f \in S_k(\mathfrak{c},\Gamma_1^1(\mathfrak{N});R)$ and a fractional O-ideal $\mathfrak{z}$ prime to p. Since we see from (ex3) applied to $z \in Z(\mathbb{A}^{(p\infty)})$, the factor $|\det(z)|_\mathbb{A} = |z|_\mathbb{A}^2$ in front of the summation symbol $\sum_h$ is canceled out by the factor $|z|_\mathbb{A}^{-2}$ in (4.41), we only need to write explicitly $\mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q) \otimes \mathfrak{z}^{-1}$ for $\mathfrak{z} = z\widehat{O} \cap F$ in order to compute the q-expansion at the cusp $(\mathfrak{a},\mathfrak{b})$ of $f|\langle\mathfrak{z}\rangle$. Since Y_z in (4.45) is given by $A \otimes \mathfrak{z}^{-1} = A/A[\mathfrak{z}]$, for an integral idele z with $\mathfrak{z} = F \cap z\widehat{O}$, we find

$$\mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q)\otimes\mathfrak{z}^{-1}=\frac{\widehat{\mathbb{G}}_m\otimes\mathfrak{a}^*/q^{\mathfrak{b}}}{q^{\mathfrak{z}^{-1}\mathfrak{b}}/q^{\mathfrak{b}}\times\mu_{\mathfrak{z}}}=\frac{\widehat{\mathbb{G}}_m\otimes\mathfrak{a}^*/q^{\mathfrak{z}^{-1}\mathfrak{b}}}{\mu_{\mathfrak{z}}}.$$

Tensoring $\widehat{\mathbb{G}}_m$ with the exact sequence: $0\to\mathfrak{a}^*\to(\mathfrak{z}\mathfrak{a})^*\to O/\mathfrak{z}\to 0$, we get

$$0\to(\mu_{\mathfrak{z}}\cong)\,\mathrm{Tor}_1(O/\mathfrak{z},\widehat{\mathbb{G}}_m)\to\widehat{\mathbb{G}}_m\otimes\mathfrak{a}^*\to\widehat{\mathbb{G}}_m\otimes(\mathfrak{z}\mathfrak{a})^*\to 0.$$

This shows

$$\mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q)\otimes\mathfrak{z}^{-1}=\frac{\widehat{\mathbb{G}}_m\otimes\mathfrak{a}^*/q^{\mathfrak{z}^{-1}\mathfrak{b}}}{\mu_{\mathfrak{z}}}\cong\frac{\widehat{\mathbb{G}}_m\otimes(\mathfrak{z}\mathfrak{a})^*}{q^{\mathfrak{z}^{-1}\mathfrak{b}}}=\mathrm{Tate}_{\mathfrak{z}\mathfrak{a},\mathfrak{z}^{-1}\mathfrak{b}}(q),$$

which implies

$$a_{\mathfrak{z}^{-1}\mathfrak{a},\mathfrak{z}\mathfrak{b}}(\xi,f|\langle\mathfrak{z}\rangle)=a_{\mathfrak{a},\mathfrak{b}}(\xi,f).\tag{4.53}$$

For an ideal $\mathfrak{z}$ prime to p generated by a totally positive element α with $\alpha\equiv 1 \mod N$, we can identify $\mathrm{Tate}_{\mathfrak{z}^{-1}\mathfrak{a},\mathfrak{z}\mathfrak{b}}(q)$ with $\mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q)$ by $t\mapsto t^\alpha$ for the variable $t\in\widehat{\mathbb{G}}_m\otimes\mathfrak{a}^*$ composed with variable change $q\mapsto q^{\alpha^{-1}}$. The variable change $q\mapsto q^{\alpha^{-1}}$ is introduced to neutralize the effect of the endomorphism $t\mapsto t^\alpha$ on the period q of $\mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q)$. When $(\mathfrak{a},\mathfrak{b})=(O,\mathfrak{c}^{-1})$, this induces an equivalence up to isogenies,

$$(\mathrm{Tate}_{\mathfrak{z}^{-1},\mathfrak{z}\mathfrak{c}^{-1}}(q),\overline{\lambda}_{can},\overline{\eta}_{can},\frac{dt}{t})\sim(\mathrm{Tate}_{O,\mathfrak{c}^{-1}}(q),\overline{\lambda}_{can},\overline{\eta}_{can}\circ\alpha^{-1},\alpha^{-1}\frac{dt}{t}),$$

where λ_{can} is the canonical polarization on $\mathrm{Tate}_{O,\mathfrak{c}^{-1}}(q)$ induced by

$$(\mathfrak{d}^{-1}\oplus\mathfrak{c}^{-1})\wedge(\mathfrak{d}^{-1}\oplus\mathfrak{c}^{-1})\cong\mathfrak{c}^*$$

sending $(x,y)\wedge(x',y')$ to $x'y-xy'$, and η_{can} is the projective limit of the level $\Gamma(N)$-structure

$$(\mathfrak{d}^{-1}\otimes\mathbb{Z}/N\mathbb{Z})\oplus\frac{\mathfrak{c}^{-1}}{N\mathfrak{c}^{-1}}\ni(\xi\otimes a,b)\mapsto\left(\exp\left(\frac{-a}{N}\right)\otimes\xi,(q^{b/N}\mod q^{\mathfrak{c}^{-1}})\right)$$

and $\overline{\eta}_{can}=\eta_{can}\circ S_0(\mathfrak{N})$. Here we choose N so that $\mathfrak{N}\cap\mathbb{Z}=(N)$. We write $\langle z\rangle_v$ $(z\in Z(\mathbb{A}))$ for the center action as in (4.45) which is independent of v if $z_p=1$. By the above computation applied to $\mathfrak{z}=(\alpha)$, writing $\alpha^{(v\infty)}$ for $\alpha^{(p\infty)}$ or $\alpha^{(\infty)}$ accordingly as $v=p$ or ∞, we have, if $\underline{\omega}^\kappa=\underline{\omega}^{mI}$ for an integer m, $a_{\mathfrak{c}^{-1}}(\xi,f|\langle\alpha^{(v\infty)}\rangle_v)=\alpha^kN(\alpha)^{-2}a_{\mathfrak{c}^{-1}}(\xi,f)$ for $\alpha\in F_+^\times$ outside v. Here the factor α^k in the right-hand side is the result of

$$\begin{aligned}f(&\mathrm{Tate}_{(\alpha)^{-1},(\alpha)\mathfrak{c}^{-1}}(q),\lambda_{can},\eta_{can},\frac{dt}{t})\\&=f(\mathrm{Tate}_{O,\mathfrak{c}^{-1}}(q),\alpha^{-2}\lambda_{can},\eta_{can}\circ\alpha^{-1},\alpha^{-1}\frac{dt}{t})\\&=\alpha^kf(\mathrm{Tate}_{O,\mathfrak{c}^{-1}}(q),\alpha^{-2}\lambda_{can},\eta_{can},\frac{dt}{t}),\end{aligned}$$

and the factor $N(\alpha)^{-2}$ comes from the shift by $|\det(\alpha^{(v\infty)})|_{\mathbb{A}}$ in (4.45) of the action of $\langle\alpha^{(v\infty)}\rangle_v$. For general κ, the action of $\alpha_{w,\varepsilon,v}$ is trivial,

$$\alpha_{w,\varepsilon,v}\omega^{\otimes k} = |\alpha^{(v\infty)}|^2\alpha_v^{2I}\varepsilon_+(\alpha^{(v\infty)})^{-1}\alpha_v^{-\kappa_1-\kappa_2+I}\omega^{\otimes k} = \omega^{\otimes k},$$

and hence $\alpha_{w,\varepsilon,v}^{(v\infty)}\omega^{\otimes k} = \varepsilon_+(\alpha^{(v\infty)})^{-1}\omega^{\otimes k}$. We have from this (replacing α in the above formula by α^{-1})

$$a_{\mathfrak{c}^{-1}}(\xi, f|\langle\alpha^{(v\infty)}\rangle_v) = \varepsilon_+(\alpha^{(v\infty)})a_{\mathfrak{c}^{-1}}(\xi, f) = \alpha^{\kappa_1+\kappa_2-I}a_{\mathfrak{c}^{-1}}(\xi, f). \tag{4.54}$$

For the sheaf $\underline{\omega}_\varepsilon^\kappa$ with general κ and ε, since we have the central character ε_+, for more general $z \in Z(\mathbb{A}^{(\infty)})$ with $z_v = 1$, we find from (4.45),

$$a_{\mathfrak{c}^{-1}}(\xi, f|\langle z\rangle_v) = \varepsilon_+(z)a_{\mathfrak{c}^{-1}}(\xi, f). \tag{4.55}$$

This is the reason why we call ε_+ the central character. Also the two identities (4.54) and (4.55) for $z = \alpha^{(v)}$ coincide.

Although we do not need this, we now insert the computation of the effect of $U(\mathfrak{p})$ for $\underline{\omega}^{mI}$ on the q-expansion, using (4.17), because $\underline{\omega}^{mI}$ extends naively to $Sh_{\overline{K}}$ giving rise to $\underline{\omega}_\varepsilon^\kappa$ for $\kappa = (0, (m-1)I)$, $\varepsilon_2 = \mathbf{1}$, and $\varepsilon_+(x) = |x|_{\mathbb{A}}^{2-m}$. In this special case, (4.16) yields the correct result. By definition, $\mathrm{Tate}_{O,\mathfrak{c}^{-1}}(q)[\mathfrak{p}]$ is isomorphic to $\mu_\mathfrak{p} \times q^{(\mathfrak{p}\mathfrak{c})^{-1}/\mathfrak{c}^{-1}}$. The nonconnected subgroups over W are given by $C_\zeta = \{\zeta^m Q^m | m \in O/\mathfrak{p}\}$ for $\zeta \in \mu_\mathfrak{p}$ and a generator Q of $q^{\mathfrak{p}\mathfrak{c}^{-1}/\mathfrak{c}^{-1}}$. Thus we see $\mathrm{Tate}_{O,\mathfrak{c}^{-1}}(q)_{/W[\mu_\mathfrak{p}][[q^\xi]]_{\xi\in\mathfrak{c}^{-1}}}/C_\zeta = \mathrm{Tate}_{(\mathfrak{p}\mathfrak{c})^{-1},O}(\zeta q)_{/W[\mu_\mathfrak{p}][[q^\xi]]_{\xi\in\mathfrak{p}^{-1}\mathfrak{c}^{-1}}}$. Writing $a_{\mathfrak{p}^{-1}\mathfrak{c}^{-1}}(\xi, f)$ for the q-expansion coefficient of $f \in G_{mI}(\Gamma_1^1(\mathfrak{c}\mathfrak{p}, \mathfrak{N}); R)$ at the cusp $(\mathfrak{p}^{-1}\mathfrak{c}^{-1}, O)$, we find the q-expansion coefficient of $f|U(\mathfrak{p})$ at the cusp $(\mathfrak{c}^{-1}, O)$ is given by

$$a_{\mathfrak{c}^{-1}}(\xi, f|U(\mathfrak{p})) = N(\mathfrak{p})^{-1}\left(\sum_{\zeta\in\mu_\mathfrak{p}} \zeta^\xi\right) a_{\mathfrak{p}^{-1}\mathfrak{c}^{-1}}(\xi, f) \quad \text{for } \xi \in \mathfrak{p}^{-1}\mathfrak{c}^{-1}.$$

By the orthogonality relation of characters $\zeta \mapsto \zeta^\xi$ of $\mu_\mathfrak{p}$, we find that only the coefficient of q^ξ for $\xi \in \mathfrak{c}^{-1}$ (not for $\xi \in \mathfrak{p}^{-1}\mathfrak{c}^{-1} - \mathfrak{c}^{-1}$) survives and is equal to $a_{\mathfrak{p}^{-1}\mathfrak{c}^{-1}}(\xi, f)$. Thus, the division by $N(\mathfrak{p})$ is canceled if R is p-torsion-free, and the operator $U(\mathfrak{p})$ is well-defined on $G_{mI}(\Gamma_1^1(\mathfrak{c}\mathfrak{p}, \mathfrak{N}); R)$ even for R with nontrivial p-torsion if $G_{mI}(\Gamma_1^1(\mathfrak{c}\mathfrak{p}, \mathfrak{N}); R) = G_{mI}(\Gamma_1^1(\mathfrak{c}\mathfrak{p}, \mathfrak{N}); W) \otimes_W R$. In particular, if $m \gg 0$ so that $\underline{\omega}^{mI}$ is very ample, $U(\mathfrak{p})$ is well-defined over R. Since the ordinary locus S_m^* of the minimal compactification $M_m^* = M^*(\mathfrak{c}, \Gamma_1^1(\mathfrak{N}))$ is affine, by Koecher's principle, we have

$$\begin{aligned} H^0(S_{\infty/R}^\circ, \underline{\omega}^{mI}) &= H^0(S_{\infty/R}^*, \underline{\omega}^{mI}) \\ &= H^0(S_{\infty/W}^*, \underline{\omega}^{mI}) \otimes_W R = H^0(S_{\infty/W}^\circ, \underline{\omega}^{mI}) \otimes_W R. \end{aligned}$$

Thus the Hecke operator $U(\mathfrak{p})$ is well-defined on the space of p-adic modular forms with coefficients in any p-adically complete W-algebras R.

Suppose that $\mathfrak{q}$ is a prime factor of $(p) \cap \mathfrak{N}$. Let ϖ be a generator of $\mathfrak{q}$ with $\varpi \in O_{(v)+}^{\times}$. We compute $U(\mathfrak{q})$ for $\underline{\omega}^{mI}$ using (4.16). Since we identify each $\mathfrak{c}$-polarization with a $\varpi\mathfrak{c}$-polarization by $\lambda \mapsto \varpi^{-1}\lambda$, we are identifying the q-expansion at the cusp $(O, \mathfrak{c}^{-1})$ with that at $(O, \varpi^{-1}\mathfrak{c}^{-1})$ by inclusion $R[[q^{\xi}]]_{\xi \in \mathfrak{c}^{-1}} \subset R[[q^{\varpi^{-1}\xi}]]_{\xi \in \mathfrak{c}^{-1}}$. Since the identification $\lambda \mapsto \varpi^{-1}\lambda$ is only canonical up to O-units, the operator $U(\mathfrak{q})$ thus defined as an endomorphism of $G_{mI}(\mathfrak{c}, \Gamma; R)$ (for $\Gamma = \Gamma_0(\mathfrak{N}), \Gamma_1(\mathfrak{N})$, and $\Gamma_1^1(\mathfrak{N})$) depends on the choice of ϖ; so, we wrote it as $U(\varpi)$. Then by the same computations as above, we have, if ϖ is a global generator of $\mathfrak{q}$, $a_{\mathfrak{c}^{-1}}(\xi, f|U(\varpi)) = a_{\mathfrak{c}^{-1}}(\varpi\xi, f)$. A slightly more generally, if a totally positive, $\varpi \in F$ generates $\mathfrak{p}^h$ for an integer $h > 0$, the operator $U(\varpi) = U(\mathfrak{p})^h$ is intrinsically defined on $\underline{\omega}^k$ for general k (not necessarily equal to mI). The above computation can be performed for general k if $\mathfrak{p}^h = (\varpi)$ and yields the same result

$$a_{\mathfrak{c}^{-1}}(\xi, f|U(\varpi)) = a_{\mathfrak{c}^{-1}}(\varpi\xi, f). \tag{4.56}$$

In general, we have

$$a_{\mathfrak{c}^{-1}}(\xi, f|U(\mathfrak{q})) = a_{\mathfrak{q}^{-1}\mathfrak{c}^{-1}}(\xi, f) \text{ if } \mathfrak{q}|p\mathfrak{N} \text{ is a prime.} \tag{4.57}$$

Now we generalize the above formula to $\underline{\omega}_{\varepsilon}^{\kappa}$ using (4.44). Recall the prime element $\varpi_{\mathfrak{q}}$ we have chosen in (ex0) so that $\varepsilon_2(\varpi_{\mathfrak{q}}) = \varpi_{\mathfrak{q}}^{\kappa_1} = 1$. Then by (4.41), we find that $\varepsilon_{\Delta}^{v}\left(\begin{smallmatrix} 1 & u \\ 0 & \varpi_{\mathfrak{q}} \end{smallmatrix}\right) = 1$. If $\varpi = \varpi_{\mathfrak{q}}$, then by (4.44), the definition of $U_v(\varpi_{\mathfrak{q}})$ is identical to (4.16). Then the same computation for the sheaf $\underline{\omega}_{\varepsilon}^{\kappa}$ yields

$$a_{\mathfrak{c}^{-1}}(\xi, f|U_v(\varpi_{\mathfrak{q}})) = a_{\mathfrak{q}^{-1}\mathfrak{c}^{-1}}(\xi, f).$$

For a general local generator $\varpi'_{\mathfrak{q}}$ of $\mathfrak{q}O_{\mathfrak{q}}$, we have

$$a_{\mathfrak{c}^{-1}}(\xi, f|U_v(\varpi'_{\mathfrak{q}})) = \varpi'^{-\kappa_1}_{\mathfrak{q},v}\varepsilon_2(\varpi'_{\mathfrak{q}})a_{\mathfrak{q}^{-1}\mathfrak{c}^{-1}}(\xi, f) \text{ if } \mathfrak{q}|p\mathfrak{N} \text{ is a prime,} \tag{4.58}$$

where $\varpi'^{-\kappa_1}_{\mathfrak{q},v}\varepsilon_2(\varpi'_{\mathfrak{q}})$ is always a p-adic unit by our way of extending ε_2 to $T(\mathbb{A}^{(\infty)})$ (see (ex1) in 4.2.6) and the definition of $\varpi'^{-\kappa_1}_{\mathfrak{q}}$ in (ex0).

The formulas (4.56) and (4.57) are the special case of (4.58) for $\varepsilon = \mathbf{1}$ and $\kappa_1 = 0$. This finishes the proof for the assertion under (h2), because we have now established the W-integrality of $f|U_p(\varpi_{\mathfrak{p}})$ for all prime factors $\mathfrak{p}|p$ by Corollary 4.23 if f is W-integral. Hence the assertion under (h2) for $R = W_m$ follows from (h1). Since D_{ℓ} is dense in the space of p-adic modular forms (Theorems 4.9 and 4.10), a global W_m-integral section f of $\underline{\omega}_{\varepsilon}^{\kappa}$ is a reduction modulo p^m of a W-integral section in $D_{\ell/W}$ for a sufficiently large ℓ. By the stability of $G_{\kappa}(K, \varepsilon; W \otimes_{\mathbb{Z}} \mathbb{Q})$ under the Hecke operators, D_{ℓ} is stable under $[KgK]_p$ for $g \in \Delta_0(p^r\mathfrak{N})$, and we have proven that the space of p-adic modular forms $V(\Gamma; W)$ in (4.20) is also stable under the Hecke operators.

Under (h3), R is W-flat, and $G_{\kappa}(K, \varepsilon; R) = G_{\kappa}(K, \varepsilon; W) \otimes_W R$ (flat base change). Thus the assertion under (h3) follows from the assertion under (h2) already proved.

Although we do not need this, we again insert the computation of the effect of $T(\varpi) := T(1,\varpi)$ for $\underline{\omega}^{mI}$ on the q-expansion, using (4.17). For a prime ideal $\mathfrak{q}$ outside $\mathfrak{N}$, we have one more (order $\mathfrak{q}$) étale cyclic subgroup $\mu_\mathfrak{q}$ of $\mathrm{Tate}_{O,\mathfrak{c}^{-1}}(q)$ besides $C_\zeta = \{\zeta^m Q^m \in \mathrm{Tate}_{O,\mathfrak{c}^{-1}}(q) | m \in O\}$ for a generator Q of $q^{\mathfrak{q}^{-1}\mathfrak{c}^{-1}/\mathfrak{c}^{-1}}$. Consider the $\mathfrak{q}$-power map $\mathrm{Tate}_{O,\mathfrak{c}^{-1}}(q) \xrightarrow{\pi_\mathfrak{q}} \mathrm{Tate}_{O,\mathfrak{q}\mathfrak{c}^{-1}}(q)$ with kernel $\mu_\mathfrak{q}$. If $\mathfrak{q} = \varpi O$ ($\varpi \in F_+^\times$), this map is realized by $t \mapsto t^\varpi$. Then $\omega_{can}(q) = \frac{dt}{t}$ on $\mathrm{Tate}_{O,\mathfrak{c}^{-1}}(q)$ is sent to $(\pi_\mathfrak{q}^*)^{-1}\omega_{can}(q) = \varpi^{-1}\omega_{can}(q^\varpi)$. Since we identify $\mathfrak{c}$-polarization with $\varpi^{-1}\mathfrak{c}$-polarization by $\lambda \mapsto \varpi\lambda$, we are identifying the q-expansion at the cusp $(O,\mathfrak{c}^{-1})$ with that of $(O,\varpi\mathfrak{c}^{-1})$ by the inclusion $R[[q^{\varpi\xi}]]_{\xi\in\mathfrak{c}^{-1}} \subset R[[q^\xi]]_{\xi\in\mathfrak{c}^{-1}}$; so, we find, if $\mathfrak{q}$ is a prime ideal outside $\mathfrak{N}$ generated by a totally positive $\varpi \in F$,

$$a_{\mathfrak{c}^{-1}}(\xi, f|T(\varpi)) = a_{\mathfrak{c}^{-1}}(\varpi\xi, f) + \varpi^{mI}N(\mathfrak{q})^{-1}a_{\mathfrak{c}^{-1}}(\frac{\xi}{\varpi}, f), \tag{4.59}$$

where, as a convention, we have $a_{\mathfrak{c}^{-1}}(\xi, f) = 0$ if $\xi \notin \mathfrak{c}^{-1}$. Here the ϖ^{mI} in front of the second term comes from

$$f(\mathrm{Tate}_{O,\mathfrak{q}\mathfrak{c}^{-1}}(q), \varpi\lambda, \varpi^{-1}\omega_{can}(q^\varpi)) = \varpi^{mI} f(\mathrm{Tate}_{O,\mathfrak{q}\mathfrak{c}^{-1}}(q), \varpi\lambda, \omega_{can}(q^\varpi)).$$

For a prime ideal $\mathfrak{q}$ outside $(v)\cap\mathfrak{N}$, we now perform a similar computation for the sheaf $\underline{\omega}_\varepsilon^\kappa$ on which $\Delta_0(\mathfrak{N})$ acts. In this setting, $T_\infty(1,\varpi_\mathfrak{q}) = T_p(1,\varpi_\mathfrak{q})$; so, we may assume that $v = \infty$. We implicitly suppose all order $\mathfrak{q}$ subgroups of test objects are étale (though after computing q-expansion, we will know in many cases that the formula is valid even if it is not the case). The central element $z \in Z(\mathbb{A})$ acts by the character $z \mapsto |z^{(\infty)}|_\mathbb{A}^{-2} z_v^{-2I} \varepsilon_+^v(z^{(\infty)})$ (see (4.41)). By the decomposition $K_\mathfrak{q}\left(\begin{smallmatrix}1 & 0\\ 0 & \varpi_\mathfrak{q}\end{smallmatrix}\right)K_\mathfrak{q} = K_\mathfrak{q}\left(\begin{smallmatrix}\varpi_\mathfrak{q} & 0\\ 0 & 1\end{smallmatrix}\right) \sqcup \left(\bigsqcup_{u \bmod \mathfrak{q}} K_\mathfrak{q}\left(\begin{smallmatrix}1 & u\\ 0 & \varpi_\mathfrak{q}\end{smallmatrix}\right)\right)$, the latter half has the same effect as $\varepsilon_\Delta^\infty\left(\begin{smallmatrix}1 & u\\ 0 & \varpi_\mathfrak{q}\end{smallmatrix}\right)U(\varpi_\mathfrak{q})$ for $\varepsilon_\Delta^\infty\left(\begin{smallmatrix}1 & u\\ 0 & \varpi_\mathfrak{q}\end{smallmatrix}\right) = \varepsilon_2(\varpi_\mathfrak{q}) = 1$; so, we need to compute the effect of the first coset $K_\mathfrak{q}\left(\begin{smallmatrix}\varpi_\mathfrak{q} & 0\\ 0 & 1\end{smallmatrix}\right)$. Write h for $\left(\begin{smallmatrix}\varpi_\mathfrak{q} & 0\\ 0 & 1\end{smallmatrix}\right)$. Then $h = \varpi_\mathfrak{q} h'$ for $h' = \left(\begin{smallmatrix}1 & 0\\ 0 & \varpi_\mathfrak{q}^{-1}\end{smallmatrix}\right)$. By (4.41), $\varepsilon_\Delta^\infty(h') = 1$. The central action of $\varpi_\mathfrak{q}$ gives rise to the multiplication by $N(\mathfrak{q})^2\varepsilon_+(\varpi_\mathfrak{q})$, because $\underline{\omega}_{\varepsilon,\overline{K}}^\kappa$ is the quotient of $\underline{\omega}^k$ by the action $\omega^{\otimes k} \mapsto g_{w,\varepsilon,v}\omega^{\otimes k}$. Thus $|\det(h)|_\mathbb{A} N(\mathfrak{q})^2\varepsilon_+(\varpi_\mathfrak{q})\varepsilon_\Delta^\infty(h')$ is integral if $\kappa_1 + \kappa_2 \geq 0$. The effect of h' can be computed by the étale isogeny,

$$\alpha : \mathrm{Tate}_{O,\mathfrak{q}\mathfrak{c}^{-1}}(q) = (\widehat{\mathbb{G}}_m \otimes \mathfrak{d}^{-1})/q^{\mathfrak{q}\mathfrak{c}^{-1}} \twoheadrightarrow (\widehat{\mathbb{G}}_m \otimes \mathfrak{d}^{-1})/q^{\mathfrak{c}^{-1}} = \mathrm{Tate}_{O,\mathfrak{c}^{-1}}(q)$$

sending $t \mod q^{\mathfrak{q}\mathfrak{c}^{-1}}$ to $t \mod q^{\mathfrak{c}^{-1}}$. By the equivalence up to isogenies:

$$(\mathrm{Tate}_{O,\mathfrak{q}\mathfrak{c}^{-1}}(q), \overline{\lambda}_{can}, \eta_{can}, \frac{dt}{t}) \sim (\mathrm{Tate}_{O,\mathfrak{c}^{-1}}(q), \overline{\lambda}_{can}, \eta_{can}\circ h'^{-1}, \frac{dt}{t}),$$

we get, if $\mathfrak{q} \nmid \mathfrak{N}$ is a prime,

$$a_{\mathfrak{c}^{-1}}(\xi, f|T_v(1,\varpi_\mathfrak{q})) = a_{\mathfrak{q}^{-1}\mathfrak{c}^{-1}}(\xi, f) + N(\mathfrak{q})\varepsilon_+(\varpi_\mathfrak{q})a_{\mathfrak{q}\mathfrak{c}^{-1}}(\xi, f), \tag{4.60}$$

because $\varpi_{\mathfrak{q},v} = 1$ if $\mathfrak{q}$ is prime to $(v) \cap \mathfrak{N}$. The sheaves $\underline{\omega}^k$ and $\underline{\omega}^\kappa_\varepsilon$ have the same pullback map g^* for $g \in \Delta_0(\mathfrak{N})$ only when $k = mI$ and $(\varepsilon_1^\infty, \varepsilon_2^\infty, \varepsilon_+^\infty) = (|\cdot|_{\mathbb{A}}^{-m}, \mathbf{1}, |\cdot|_{\mathbb{A}}^{m-2})$. In this special case, it is easy to see that the formulas (4.60) and (4.59) match.

We assume (h5). Since we have $[KgK]_\infty = [KgK]_p$ if $g_p = 1$, the assertion for $[KgK]$ with $g_p = 1$ follows from the argument for (h2). Thus we only need to take care of $T_\infty(1, \varpi_\mathfrak{p})$ for $\mathfrak{p}|p$. We see that $N(\mathfrak{p})\varepsilon_+(\varpi_\mathfrak{p})$ is p-integral as long as $\kappa_1 + \kappa_2 \geq 0$ and that $N(\mathfrak{p})\varepsilon_+(\varpi_\mathfrak{p}) \equiv 0 \mod p$ as long as $\kappa_1 + \kappa_2 \geq I$, because ε_+ is the finite part of an arithmetic Hecke character of infinity type $-\kappa_1 - \kappa_2 + I$. The operator $T_\infty(1, \varpi_\mathfrak{p})$ is originally defined only over $\mathbb{Q}$. However by the above formula, if $\kappa_1 + \kappa_2 \geq 0$, it sends the $\mathcal{W}$-integral subspace $G_\kappa(K, \varepsilon; \mathcal{W})$ into itself for K with $S_1^1(\mathfrak{N}) \subset K \subset S_0(\mathfrak{N})$. Since $N(\mathfrak{p})\varepsilon_+(\varpi_\mathfrak{p}) \equiv 0 \mod p$ if $\kappa_1 + \kappa_2 \geq I$, $T_\infty(1, \varpi_\mathfrak{p}) \equiv U_p(\varpi_\mathfrak{p}) \mod p$ by the q-expansion principle (Corollary 4.23 (1)). Thus for K with $S_1^1(\mathfrak{N}) \subset K \subset S_0(\mathfrak{N})$, if $G_\kappa(K, \varepsilon; W) \otimes_W R = G_\kappa(K, \varepsilon; R)$ and $\kappa_1 + \kappa_2 \geq 0$ (under unramifiedness of p in $F/\mathbb{Q}$), the operator $T_\infty(1, \varpi_\mathfrak{p})$ is actually well-defined on $G_\kappa(K, \varepsilon; R)$, and $T_\infty(1, \varpi_\mathfrak{p})$ is equal to $U_p(\varpi_\mathfrak{p})$ if $\kappa_1 + \kappa_2 \geq I$ and $pR = 0$. □

We record what we have seen for p-adic modular forms in the above proof.

Corollary 4.29 *Let Γ be a group with $\widehat{\Gamma}_1^1(\mathfrak{N}) \subset \Gamma \subset \widehat{\Gamma}_0(\mathfrak{N})$. For each p-adic W-algebra R, the space $V(\Gamma; R)$ of p-adic modular forms defined in* (4.20) *is stable under the Hecke operators $T_p(1, y)$ with $y_p = 1$, $\langle y \rangle_p$ for $y \in \widehat{O} \cap (F_\mathbb{A}^{(p\infty)})^\times$ and $U_p(y)$ for $y \in F_p^\times \cap O_p$.*

By the universality of $V(\mathfrak{c}, \Gamma; W)$ (which classifies $(A, \overline{\lambda}, \overline{\phi}, i_p)_{/R}$ for p-adic W-algebras R), we find $V(\mathfrak{c}, \Gamma; R) = V(\mathfrak{c}, \Gamma; W)\widehat{\otimes}_W R$; so, the assertion for R follows from the same assertion for W.

4.2.10 Adelic q-Expansion

In this subsection, we introduce adelic q-expansion coefficients $\mathbf{a}_p(y, f)$ for a modular form f on Sh_K, whose Archimedean counterpart is the Fourier coefficients of a modular form on $G(\mathbb{A})$ with respect to the canonical Whittaker function (whose Mellin transform is the Γ-factor of the standard L-function of $GL(2)$). Let $\mathcal{B}$ be the base p-adic W-algebra. The coefficients $\mathbf{a}_p(y, f)$ for ideles y are useful in the following section to describe the duality between the Hecke algebra $h_\kappa(\mathfrak{N}, \varepsilon; \mathcal{B})$ and the space of modular forms $S_\kappa(\mathfrak{N}, \varepsilon; \mathcal{B})$.

Choose a complete representative set $\{c\} \subset (F_\mathbb{A}^{(p\infty)})^\times$ for the class group $Cl_F^+ = F_\mathbb{A}^\times / \widehat{O}^\times F^\times F_{\infty+}^\times$ as in Section 4.1.3, and write $\mathfrak{c} = c\widehat{O} \cap F$. Recall that $Sh_{S_0(\mathfrak{N})} = \bigsqcup_{\mathfrak{c} \in Cl_F^+} \mathfrak{M}(\mathfrak{c}, \Gamma_0(\mathfrak{N}))$ with

$$\mathfrak{M}(\mathfrak{c}, \Gamma_0(\mathfrak{N}))(\mathbb{C}) = G(\mathbb{Q})_+ \backslash G(\mathbb{Q})_+ \widetilde{c} S_0(\mathfrak{N}) G(\mathbb{R})^+ / S_0(\mathfrak{N}) C_\mathbf{0},$$

where $\widetilde{c} = \left(\begin{smallmatrix} c & 0 \\ 0 & 1 \end{smallmatrix}\right)$. A modular form $f \in S_\kappa(S_0(\mathfrak{N}), \varepsilon; \mathcal{B})$ can be considered as a tuple $(f_\mathfrak{c})_{\mathfrak{c} \in Cl_F^+}$ with $f_\mathfrak{c}(A, \overline{\lambda}, \overline{\eta}, \omega)\omega^{\otimes k} \in H^0(\mathfrak{M}(\mathfrak{c}, \Gamma_0(\mathfrak{N})), \underline{\omega}_{\kappa, \varepsilon/\mathcal{B}})$ for $k =$

$\kappa_2 - \kappa_1 + I$. A p-adic modular form f is a tuple $(f_{\mathfrak{c}})_{\mathfrak{c}}$ indexed by $\mathfrak{c} \in Cl_F^+$ of sections $f_{\mathfrak{c}} \in V_{cusp}(\mathfrak{c}, \Gamma_0(\mathfrak{N}); \mathcal{B})$ of the structure sheaf of the formal completion of $S = \mathfrak{M}(\mathfrak{c}, \Gamma_0(\mathfrak{N}))[\frac{1}{E}]$ (for a lift E of the Hasse invariant) along its modulo p fiber S_1. For a given modular form $f = (f_{\mathfrak{c}})$, it has a q-expansion at the cusp $(O, \mathfrak{c}^{-1})$: $f_{\mathfrak{c}} = \sum_{0 \ll \xi \in \mathfrak{c}^{-1}} a_{\mathfrak{c}^{-1}}(\xi; f_{\mathfrak{c}}) q^{\xi}$. We want to compute the action of the matrix $g = \left(\begin{smallmatrix} \epsilon^{(\infty)} & 0 \\ 0 & 1 \end{smallmatrix}\right)$ for a totally positive unit $\epsilon \in O$ on the q-expansion of f. More generally, we define $f|g$ for $f \in G_\kappa(S(N), \varepsilon; \mathcal{B})$ and $g \in \Delta$ by

$$(f|g)(A, \overline{\lambda}, \overline{\eta}^{(p)}, \omega)\omega^{\otimes k} = g^*(f(A, \overline{\lambda}, \overline{\eta}^{(p)} \circ g^{-1}, \omega)\omega^{\otimes k}), \tag{4.61}$$

where $\omega^{\otimes k} \mapsto g^*\omega^{\otimes k}$ is as in (4.42). This is a right action. Taking $v = p$ and $\Delta = \Delta(p^\infty)_p \times G(\mathbb{A}^{(p)})$, $f \mapsto |\det(g)|_{\mathbb{A}}^{-1}(f|g)$ gives rise to a right automorphic representation of the group $\Delta(p^\infty)_p^\times \times G(\mathbb{A}^{(p)})$ on $\varinjlim_N H^0(Sh^{(p)}_{\overline{S}(N)}, \underline{\omega}^\kappa_{\varepsilon/\mathcal{B}})$. Even if $\eta^{(p)}$ misses its p-component, g_p acts nontrivially on the differential ω (so nontrivial g_p does make difference). Here N runs over all positive integers prime to p. If $\mathcal{B}$ is a $\mathbb{Q}$-algebra, taking $v = \infty$ and $\Delta = G(\mathbb{A}^{(\infty)})$, we can extend the action $f \mapsto |\det(g)|_{\mathbb{A}}(f|g)$ to $G(\mathbb{A}^{(\infty)})$ in exactly the same way, replacing $(Sh^{(p)}, \eta^{(p)})$ by (Sh, η) and taking the limit over all positive integers N. Note that the choice of $v = p$ and $v = \infty$ makes the resulting representations different at their p-component. Over $\mathbb{C}$, for the choice $v = \infty$, this is the right regular automorphic representation of $G(\mathbb{A}^{(\infty)})$ (with central character ε_+) on the space of holomorphic Hilbert modular forms and is known to be a discrete direct sum of irreducible admissible (automorphic) representations with multiplicity at most 1 ([AAG]). The twist $\otimes|\det(g)|_{\mathbb{A}}$ has to be there to undo the reverse twist $|z|_{\mathbb{A}}^{-2}$ appearing in the definition (ex3) of $\varepsilon_S^v(z)$. We have done such a twist in order to formulate the effect of Hecke operators in (4.58) and (4.60) in a traditional form. Without the twist, the right-hand side of (4.58) and (4.60) for $v = \infty$ will be divided by $N(\mathfrak{q})$. In any case, we also obtained a rational structure over $\mathbb{Q}[\varepsilon_+]$ on the (right regular) automorphic representation acting on holomorphic modular forms.

We compute the q-expansion of $f|g$ for $g = \left(\begin{smallmatrix} \epsilon^{(\infty)} & 0 \\ 0 & 1 \end{smallmatrix}\right)$ for $\epsilon \in O^\times$. Take $f \in G_\kappa(S(\mathfrak{N}), \varepsilon; \mathcal{B})$. We consider the effect of $t \mapsto t^\epsilon$ on the Tate AVRM $\mathrm{Tate}_{O,\mathfrak{c}^{-1}}(q)$. This map is induced by multiplication by ϵ on $\mathbb{G}_m \otimes \mathfrak{d}^{-1}$, which induces multiplication by ϵ on $\mu_{\mathfrak{N}}$ (corresponding to the action $i \mapsto i \circ \epsilon$ on level $\Gamma_1(\mathfrak{N})$-structure i). Since the level $\Gamma_1(\mathfrak{N})$-structure i gives the identification of $\left(\mathfrak{N}^* \oplus \mathfrak{c}^{-1}\right)/L$ with $\mu_{\mathfrak{N}} \subset \mathbb{G}_m \otimes \mathfrak{d}^{-1}$ (for $L = \mathfrak{d}^{-1} \oplus \mathfrak{c}^{-1}$), the action $t \mapsto t^\epsilon$ on level $\Gamma_1(\mathfrak{N})$-structures corresponds to the action of g^{-1} for the matrix $g = \left(\begin{smallmatrix} \epsilon^{(\infty)} & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Indeed, the isogeny $\alpha_g : A \to Y_g$ for $g \in \Delta_0(N)$ is just the quotient map $A \to Y_g = A/\operatorname{Ker}(\alpha_g)$ ($\operatorname{Ker}(\alpha_g) = \eta(g^{-1}\widehat{L}/\widehat{L})$); so, it does not have any effect on $\widehat{L}$. The action $t \mapsto t^\epsilon$ has an effect on $\mathcal{T}(\mathrm{Tate}_{O,\mathfrak{c}^{-1}}(q))$ which is canceled out by the action of g^{-1} on $\mathcal{T}(\mathrm{Tate}_{O,\mathfrak{c}^{-1}}(q))$. Since α_g and g act trivially on $q^{\mathfrak{N}^{-1}\mathfrak{c}^{-1}}/q^{\mathfrak{c}^{-1}}$, to describe the action of g on $\mathrm{Tate}_{O,\mathfrak{c}^{-1}}(q)$, we need to make variable change $q \mapsto q^{\epsilon^{-1}}$, to compensate the effect on $q^{\mathfrak{N}^{-1}\mathfrak{c}^{-1}}/q^{\mathfrak{c}^{-1}}$. The action of g is reflected by the q-expansion of $f \in S_\kappa(\mathfrak{N}, \varepsilon; \mathcal{B})$ as follows,

$a_{\mathfrak{c}^{-1}}(\xi, f|g) = a_{\mathfrak{c}^{-1}}(\epsilon\xi, f)$. Then we have

$$\begin{aligned}(f_{\mathfrak{c}}|g(A,\overline{\lambda},\overline{\eta},\omega))\omega^{\otimes k} &= g^*_{\kappa,\varepsilon}(f_{\mathfrak{c}}(A,\overline{\lambda},\overline{\eta}\circ g^{-1},\omega)\omega^{\otimes k})\\ = f_{\mathfrak{c}}(A,\overline{\lambda},\overline{\eta}\circ g^{-1},\omega)(g^*_{\kappa,\varepsilon}\omega^{\otimes k}) &= \varepsilon^p_{\Delta}(g^{(\infty)})(\det(g)_v)^w f_{\mathfrak{c}}(A,\overline{\lambda},\overline{\eta}\circ g^{-1},\omega)\omega^{\otimes k}\\ &= \varepsilon_1(\epsilon_{p\mathfrak{N}})\epsilon^{-\kappa_1} f_{\mathfrak{c}}(A,\overline{\lambda},\overline{\eta}\circ g^{-1},\omega)\omega^{\otimes k},\end{aligned}$$

where the last identity follows from $\kappa_1 = I - w$. Note that $\varepsilon_1(\epsilon_{p\mathfrak{N}}) = \varepsilon_1(\epsilon)$ by (ex1) and that $\varepsilon_+(\epsilon) = 1$ for a unit $\epsilon \in O_+^\times$. This shows

$$a_{\mathfrak{c}^{-1}}(\epsilon\xi, f) = \varepsilon_1(\epsilon)\epsilon^{-\kappa_1} a_{\mathfrak{c}^{-1}}(\xi, f) \quad \text{for } f \in S_\kappa(\mathfrak{N}, \varepsilon; \mathcal{B}). \tag{4.62}$$

In Section 4.1.9, we only considered p-adic modular forms on PG with a trivial central character. Since we have created the line bundle $\underline{\omega}^\kappa_\varepsilon$ over $Sh(PG, X)$, moving around κ fixing the central character $\varepsilon_+ : Z(\mathbb{A}^{(\infty)}) \to \mathcal{W}^\times$ and $\varepsilon_\jmath : T(\widehat{\mathbb{Z}}) \to \mathcal{W}^\times$, we have a similar theory for p-adic modular forms with a fixed central character. We now describe the theory. Write $\widehat{\Gamma}$ (resp. Γ) for $\widehat{\Gamma}_\jmath(\mathfrak{N})$ (resp. $\Gamma_\jmath(\mathfrak{N})$) for $\jmath = 0, 1$. Define $V_{cusp}(\widehat{\Gamma}, \varepsilon; \mathcal{B})$ to be the space of p-adic modular forms (with coefficients in a p-adic W-algebra $\mathcal{B}$) with central character ε_+ and Neben character $\varepsilon_\jmath$. Here we assume that the conductor of ε^- is a factor of $\mathfrak{N}$ so that ε_Δ is well-defined over $\Delta_0(\mathfrak{N})$. Then $V_{cusp}(\widehat{\Gamma}, \varepsilon; W)$ is the space of global formal sections of the sheaf $\underline{\widehat{\omega}}_{\kappa,\varepsilon} = \varprojlim_m \underline{\omega}_{\kappa,\varepsilon/S_m}$ over the Igusa tower $T_{\infty,\infty}/S_\infty$, where $S_m = Sh^{(p)}(PG, X)/S_0(N)\left[\frac{1}{E}\right]_{/W_m}$ for a lift E of the Hasse invariant. Here $Sh^{(p)}(PG, X)/S_0(N) = \bigsqcup_{\mathfrak{c}\in Cl^{[2]}} \mathfrak{M}(\mathfrak{c}, \Gamma)_{/W_m}$ for a complete representative set $\{\mathfrak{c}\}$ for the two-torsion group $Cl^{[2]} = Cl_F^+/(Cl_F^+)^2$, and physically as the $T(\mathbb{Z}_p)$-module (forgetting the Hecke operators), we have

$$V_{cusp}(\widehat{\Gamma}, \varepsilon; \mathcal{B}) = \bigoplus_{\mathfrak{c}\in Cl^{[2]}} V_{cusp}(\mathfrak{c}, \Gamma; \mathcal{B}).$$

In other words, if we identify $V_{cusp}(\mathfrak{c}\mathfrak{z}^2, \Gamma)$ with $V_{cusp}(\mathfrak{c}, \Gamma)$ by $f \leftrightarrow f|\langle\mathfrak{z}\rangle$ for the operator $\langle\mathfrak{z}\rangle$ introduced in Section 4.1.9, we get $V_{cusp}(\widehat{\Gamma}, \mathbf{1}; \mathcal{B}) = V_{cusp}(\widehat{\Gamma})$ of Section 4.1.9. Instead we identify $V_{cusp}(\mathfrak{c}\mathfrak{z}^2, \Gamma; \mathcal{B})$ with $V_{cusp}(\mathfrak{c}, \Gamma; \mathcal{B})$ by $\varepsilon_+(\mathfrak{z})^{-1}\langle\mathfrak{z}\rangle$. We still have $V_{cusp}(\widehat{\Gamma}, \varepsilon; \mathcal{B}) = \bigoplus_{\mathfrak{c}\in Cl^{[2]}} V_{cusp}(\mathfrak{c}, \Gamma; \mathcal{B})$ (under the different identification). Thus, forgetting about the action of $T(\widehat{\mathbb{Z}}^{(p)})$, as a module over $T(\mathbb{Z}_p)$, we have $V_{cusp}(\widehat{\Gamma}, \varepsilon; \mathcal{B}) = V_{cusp}(\widehat{\Gamma}; \mathcal{B})$ physically. So, Theorems 4.11 and 4.12 remain valid for $V_{cusp}(\widehat{\Gamma}, \varepsilon; W)$ with $\widehat{\Gamma} = \widehat{\Gamma}_0(\mathfrak{N})$ and $\widehat{\Gamma}_1(\mathfrak{N})$. In particular, the W-dual of $V^{ord}_{cusp}(\widehat{\Gamma}, \varepsilon; W)$ is $W[[T(\mathbb{Z}_p)]]$-projective (of finite rank) at least if $p \geq 5$.

Replacing $\overline{\eta}^{(p)}$ by the equivalent cyclic subgroup C of order $\mathfrak{N}$ to indicate that this is a level $\Gamma_0(\mathfrak{N})$-structure and denoting by $i_p : \mu_{p^\infty} \otimes \mathfrak{d}^{-1} \hookrightarrow A[p^\infty]$ for level $\Gamma_1(p^\infty)$-structures, we may think of $f \in V_{cusp}(\widehat{\Gamma}_0(\mathfrak{N}), \varepsilon; \mathcal{B})$ as a functorial rule $(A, \overline{\lambda}, C, i_p)_{/R} \mapsto f(A, \overline{\lambda}, C, i_p) \in R$ satisfying the following three properties for (variable) p-adic $\mathcal{B}$-algebras R and R'.

(G_p1) $(A,\lambda,C,i_p)_{/R} \cong (A',\lambda',C',i'_p)_{/R} \Rightarrow f(A,\lambda,C,i_p) = f(A',\lambda',C',i'_p)$;
(G_p2) If $\rho : R \to R'$ is a continuous W-algebra homomorphism, then

$$f((A,\lambda,C,i_p) \times_{R,\rho} R') = \rho(f(A,\lambda,C,i_p));$$

(Z) $f|\langle\mathfrak{z}\rangle = \varepsilon_+(\mathfrak{z})f$ for all fractional ideals $\mathfrak{z}$ prime to $p\mathfrak{N}$ ($\Leftrightarrow f|z = \varepsilon_+(z)f$ for all $z \in Z(\mathbb{A}^{(p\mathfrak{N}\infty)})$).

By Koecher's principle, we do not need to have the condition corresponding to (G_p3) in Section 3.2.9 if $F \neq \mathbb{Q}$; so, we have omitted it. We can let $a \in T(\mathbb{Z}_p)$ act on $f \in V_{cusp}(\widehat{\Gamma}_0(\mathfrak{N}),\varepsilon;\mathcal{B})$ by $f(A,\lambda,C,i_p) \mapsto f(A,\lambda,C,i_p \circ a)$. In the above definition, replacing the level $\Gamma_0(\mathfrak{N})$-structure C by a level $\Gamma_1(\mathfrak{N})$-structure $i_{\mathfrak{N}} : \mu_{\mathfrak{N}} \hookrightarrow A[\mathfrak{N}]$, we get the definition of the space of p-adic modular forms $V_{cusp}(\widehat{\Gamma}_1(\mathfrak{N}),\varepsilon;\mathcal{B})$ on $\Gamma_1(\mathfrak{N})$. Since $Sh_{S_1(\mathfrak{N})}$ is a disjoint union $\bigsqcup_{\mathfrak{c}\in Cl_F^+} \mathfrak{M}(\mathfrak{c},\Gamma_1(\mathfrak{N}))$, we have an embedding $V_{cusp}(\widehat{\Gamma}_1(\mathfrak{N}),\varepsilon;\mathcal{B})$ into $\bigoplus_{\mathfrak{c}\in Cl_F^+} V_{cusp}(\mathfrak{c},\Gamma_1(\mathfrak{N});\mathcal{B})$.

We return to the level $\Gamma_0(\mathfrak{N})$. We write $V_{cusp}(\widehat{\Gamma}_0(\mathfrak{N}),\varepsilon;\mathcal{B})[\kappa_1,\varepsilon_1]$ for the subspace on which $T(\widehat{\mathbb{Z}})$ acts by the character $w\varepsilon_1 : z \mapsto z_p^{-\kappa_1}\varepsilon_1(z)$ (note that $\kappa_1 = I - w$). We have a natural map $S_\kappa(\mathfrak{N},\varepsilon;\mathcal{B}) \hookrightarrow V_{cusp}(\widehat{\Gamma}_0(\mathfrak{N}),\varepsilon;\mathcal{B})[\kappa_1,\varepsilon_1]$ by defining $f(A,\lambda,C,i_p) := f(A,\overline{\lambda},\overline{\eta}^{(p)},i_{p,*}(\frac{dt}{t}\otimes 1))$ for the level $S_0(\mathfrak{N})$-structure $\overline{\eta}^{(p)}$ induced by C and for the differential $\frac{dt}{t}$ on $\mathbb{G}_m \otimes \mathfrak{d}^{-1}$. By the q-expansion principle Corollary 4.23, this map is injective. We verify that (4.62) still holds for $f \in V_{cusp}(\widehat{\Gamma}_0(\mathfrak{N}),\varepsilon;\mathcal{B})[\kappa_1,\varepsilon_1]$ by a similar computation.

Choose a finite idele c so that $cO = \mathfrak{c}$ and $c_p = 1$. For finite integral ideles y and $f \in V_{cusp}(\widehat{\Gamma}_1(\mathfrak{N}),\varepsilon;\mathcal{B})[\kappa_1,\varepsilon_1]$, we introduce a continuous function $y \mapsto \mathbf{a}_p(y;f) \in \mathcal{B}$ as follows: Choose ideal representatives $\mathfrak{c}$ (prime to Np) so that $yO = \xi\mathfrak{c}$ for $\xi \in \mathfrak{c}_+^{-1}$, and write $y = u\xi c$ for ideles u,c with $c_p = c_\mathfrak{l} = 1$ for all prime factors $\mathfrak{l}$ dividing the conductor of ε and $c\widehat{O} \cap F = \mathfrak{c}$. Define

$$\mathbf{a}_p(y;f) = \varepsilon_1(u_{p\mathfrak{N}})u_p^{-\kappa_1} \cdot a_{\mathfrak{c}^{-1}}(\xi;f_\mathfrak{c}) = a_{\mathfrak{c}^{-1}}(\xi;f_\mathfrak{c}|\left(\begin{smallmatrix} u & 0 \\ 0 & 1 \end{smallmatrix}\right)). \tag{4.63}$$

The definition (4.63) does not depend on the choice of c. Indeed, taking another c' and $\mathfrak{c}' = F \cap c'\widehat{O}$ with $y = u'\xi'c'$, by $\mathrm{Tate}_{O,\mathfrak{c}^{-1}}(q) \cong \mathrm{Tate}_{O,\mathfrak{c}'^{-1}}(q)$ given via $t \mapsto t^\varpi$ for $\varpi = \xi^{-1}\xi'$ (which pulls back $\frac{dt}{t}$ on $\mathrm{Tate}_{O,\mathfrak{c}'^{-1}}(q)$ to $\varpi\frac{dt}{t}$ on $\mathrm{Tate}_{O,\mathfrak{c}^{-1}}(q)$), we have $a_{\mathfrak{c}^{-1}}(\xi;f_\mathfrak{c}|\left(\begin{smallmatrix} u & 0 \\ 0 & 1 \end{smallmatrix}\right)) = a_{\mathfrak{c}'^{-1}}(\xi';f_{\mathfrak{c}'}|\left(\begin{smallmatrix} u' & 0 \\ 0 & 1 \end{smallmatrix}\right))$, because $uu'^{-1} = \varpi^{-1}$ which compensates the action of $t \mapsto t^\varpi$ on the level structure. By definition, if $f \in V_{cusp}(\widehat{\Gamma}_1(\mathfrak{N}),\varepsilon;\mathcal{B})[\kappa_1,\varepsilon_1]$, then $\mathbf{a}_p(uy;f) = u_p^{-\kappa_1}\varepsilon_1(u)\mathbf{a}_p(y;f)$ for $u \in T(\widehat{\mathbb{Z}})$. Thus $\mathbf{a}_p$ is well-defined independently of the choice of c, and for an integral ideal $\mathfrak{n}$ prime to p and the conductor $\mathfrak{C}_1$ of ε_1, choosing a finite idele n so that $nO = \mathfrak{n}$, $\mathbf{a}_p(n^{(p\mathfrak{C}_1)}y;f)$ is well-defined independently of the choice of n. We write $\mathbf{a}_p(y\mathfrak{n};f) = \mathbf{a}_p(n^{(p\mathfrak{C}_1)}y;f)$. We extend the function $\mathbf{a}_p$ outside integral ideles by defining the value to be 0.

The above definition of $\mathbf{a}_p(y,f)$ extends without modification to p-adic modular forms f in $V_{cusp}(\widehat{\Gamma}_1(\mathfrak{N}),\varepsilon;\mathcal{B})$ regarding $V_{cusp}(\widehat{\Gamma}_1(\mathfrak{N});\mathcal{B})$ as a subspace

of the direct sum $\bigoplus_{\mathfrak{c}\in Cl_F^+} V_{cusp}(\mathfrak{c}, \Gamma_1(\mathfrak{N}); \mathcal{B})$, because we have a naturally defined action of $\left(\begin{smallmatrix} u & 0 \\ 0 & 1 \end{smallmatrix}\right)$ for $u \in T(\widehat{\mathbb{Z}})$ on such p-adic modular forms f given by $f|\left(\begin{smallmatrix} u & 0 \\ 0 & 1 \end{smallmatrix}\right)(A, \lambda, i_{\mathfrak{N}}, i_p) = \varepsilon_1(u) f(A, \lambda, i_{\mathfrak{N}} \circ u_{\mathfrak{N}}^{-1}, i_p \circ u_p^{-1})$. By Corollary 4.23, the p-adic modular form is determined by the function $\mathbf{a}_p$ on integral ideles. We have now proved

$$V_{cusp}(\widehat{\Gamma}_0(\mathfrak{N}), \varepsilon; \mathcal{B})[\kappa_1, \varepsilon_1] = \left\{ f \in V_{cusp}(\widehat{\Gamma}_0(\mathfrak{N}), \varepsilon; \mathcal{B}) \middle| \mathbf{a}_p(uy, f) = \varepsilon_1(u) u_p^{-\kappa_1} \mathbf{a}_p(y, f)\ \forall u \in T(\widehat{\mathbb{Z}}) \right\}.$$

We compute the effect of $T_p(\varpi_{\mathfrak{q}}) := T_p(1, \varpi_{\mathfrak{q}})$ for a prime $\mathfrak{q} \nmid \mathfrak{N}$ on the coefficients $\mathbf{a}_p$ using (4.60). We have, if $\mathfrak{q} \nmid p\mathfrak{N}$,

$$\begin{aligned} \mathbf{a}_p(y, f|T_p(\varpi_{\mathfrak{q}})) =& a_{\mathfrak{c}^{-1}}(\xi; f_{\mathfrak{c}}|T_p(\varpi_{\mathfrak{q}})|\left(\begin{smallmatrix} u & 0 \\ 0 & 1 \end{smallmatrix}\right)) \\ =& a_{\mathfrak{q}^{-1}\mathfrak{c}^{-1}}(\xi; f_{\mathfrak{q}\mathfrak{c}}|\left(\begin{smallmatrix} u & 0 \\ 0 & 1 \end{smallmatrix}\right)) \\ &+ N(\mathfrak{q})\varepsilon_+(\varpi_{\mathfrak{q}}) a_{\mathfrak{q}\mathfrak{c}^{-1}}\left(\xi; f_{\mathfrak{q}^{-1}\mathfrak{c}}|\left(\begin{smallmatrix} u & 0 \\ 0 & 1 \end{smallmatrix}\right)\right). \end{aligned} \tag{4.64}$$

Note that $y = u\xi c$ implies $\varpi_{\mathfrak{q}} y = u\xi c \varpi_{\mathfrak{q}}$ and $\varpi_{\mathfrak{q}}^{-1} y = u\xi c \varpi_{\mathfrak{q}}^{-1}$. Then replacing c (resp. c) by $c\varpi_{\mathfrak{q}}$ (resp. $c\varpi_{\mathfrak{q}}^{-1}$) in the first (resp. the second) term of the last equation of (4.64) and applying the definition (4.63) of $\mathbf{a}_p(\varpi_{\mathfrak{q}} y, f)$ and $\mathbf{a}_p(\varpi_{\mathfrak{q}}^{-1} y, f)$, we find, for all primes $\mathfrak{q}$ outside $p\mathfrak{N}$,

$$\mathbf{a}_p(y, f|T_p(\varpi_{\mathfrak{q}})) = \mathbf{a}_p(\varpi_{\mathfrak{q}} y, f) + \varepsilon_+(\varpi_{\mathfrak{q}}) N(\mathfrak{q}) \mathbf{a}_p(\frac{y}{\varpi_{\mathfrak{q}}}, f). \tag{4.65}$$

The same formula is valid for $T_\infty(\varpi_{\mathfrak{q}}) := T_\infty(1, \varpi_{\mathfrak{q}})$ including $\mathfrak{q}|p$ for our chosen prime element $\varpi_{\mathfrak{q}}$ (because $\varepsilon_1(\varpi_{\mathfrak{q}}) = 1$); otherwise, the formula depends on $v = p, \infty$. Similarly, we have

$$\mathbf{a}_p(y, f|U_p(\varpi_{\mathfrak{p}})) = \mathbf{a}_p(\varpi_{\mathfrak{p}} y, f). \tag{4.66}$$

We write $T_v(y)$ $(v = p, \infty)$ for the Hecke operator acting on $G_\kappa(\mathfrak{N}, \varepsilon; \mathcal{B})$ corresponding to the double coset decomposition of

$$\mathcal{T}(y) = \left\{ x \in \Delta_0(\mathfrak{N}) \middle| \det(x)\widehat{O} = y\widehat{O} \right\}$$

for $y \in \widehat{O} \cap F^\times_{\mathbb{A}^{(\infty)}}$. Note that, if $yO = y\widehat{O} \cap F$ is square-free (as an ideal), we have $T_v(y) = T_v(1, y)$, but otherwise, they are different.

Since we have a relation $T(\varpi_{\mathfrak{q}})^2 = T(\varpi_{\mathfrak{q}}^2) + N(\mathfrak{q})\langle \varpi_{\mathfrak{q}} \rangle$ and $U(\varpi_{\mathfrak{q}})^n = U(\varpi_{\mathfrak{q}}^n)$ (cf. [MFG] (3.27)), by iteration, we get from (4.66) and (4.65),

$$\mathbf{a}_p(1, f|T_p(y)) = \mathbf{a}_p(y, f) \quad \text{for all integral ideles } y. \tag{4.67}$$

4.2.11 Nearly Ordinary Hecke Algebra with Central Character

We now introduce the nearly ordinary Hecke algebra and prove the duality between the algebra and the space of cusp forms.

We take limits $e = \lim_{n\to\infty} U_p(p)^{n!}$ and $e_0 = \lim_{n\to\infty} T_\infty(p)^{n!}$ in the algebra $h_\kappa(K,\varepsilon;R)$ defined in Section 4.2.8 for a profinite W-algebra R, whenever the Hecke operator $U_p(p) = U_\infty(p)$ or $T_\infty(p)$ is in $h_\kappa(K,\varepsilon;R)$. We define for K with $S_1^1(\mathfrak{N}) \subset K \subset S_0(\mathfrak{N})$,

$$\begin{aligned} S_\kappa^{n.ord}(K,\varepsilon;R) &= \begin{cases} e\,(S_\kappa(K,\varepsilon;R)) & \text{if } p|\mathfrak{N}, \\ e_0\,(S_\kappa(K,\varepsilon;R)) & \text{if } \mathfrak{N} \text{ prime to } p, \end{cases} \\ h_\kappa^{n.ord}(K,\varepsilon;R) &= \begin{cases} e\,(h_\kappa(K,\varepsilon;R)) & \text{if } p|\mathfrak{N}, \\ e_0\,(h_\kappa(K,\varepsilon;R)) & \text{if } \mathfrak{N} \text{ prime to } p. \end{cases} \end{aligned} \tag{4.68}$$

The algebra $h_\kappa^{n.ord}(K,\varepsilon;R)$ is called the nearly ordinary Hecke algebra on G of level K with Neben (or weight) character (κ,ε) and with coefficients in R. We split $T(\mathbb{Z}_p) = \Gamma \times \Delta^{(p)}$ for the maximal p-profinite subgroup Γ and the maximal prime-to-p torsion (finite) group $\Delta^{(p)}$. Since p is unramified in $F/\mathbb{Q}$, Γ is torsion-free. For each p-profinite W-algebra R, any projective $R[[T(\mathbb{Z}_p)]]$-module of finite type is $R[[\Gamma]]$-free of finite rank. We then define $\mathbf{h}^{n.ord}(\mathfrak{N},\varepsilon;W[[T(\mathbb{Z}_p)]])$ for the $W[[T(\mathbb{Z}_p)]]$-subalgebra of $\mathrm{End}(\mathcal{V}_{cusp}^{ord}(\widehat{\Gamma}_0(\mathfrak{N}),\varepsilon;R))$ generated over $W[[T(\mathbb{Z}_p)]]$ by the operators $T_p(y)$ and $\langle y\rangle_p$ for all integral ideles y.

Theorem 4.30 *Let K be a subgroup of $G(\mathbb{A}^{(\infty)})$ with $S_1(p^r)\cap S_1^1(\mathfrak{N}) \subset K \subset S_0(p^r\mathfrak{N})$ for $\mathfrak{N}$ prime to p. By the pairing $\langle t, f\rangle = \mathbf{a}_p(1, f|t)$ for modular forms f and Hecke operators t, we have the following perfect duality.*

(1) *Suppose that K is maximal at p and that κ is a dominant weight of T_G ($\Leftrightarrow \kappa_2 \geq \kappa_1$). If one of the conditions* (h4–6) *in* Theorem 4.28 *is met depending on the level, we have $h_\kappa(K,\varepsilon;R) \cong \mathrm{Hom}_R(S_\kappa(K,\varepsilon;R),R)$ and $S_\kappa(K,\varepsilon;R) \cong \mathrm{Hom}_R(h_\kappa(K,\varepsilon;R),R)$ for an algebra R finite over $\mathcal{W}$;*

(2) *If $r > 0$ and one of the conditions* (h1–3) *is met,*
$$h_\kappa^{n.ord}(K,\varepsilon;R) \cong \mathrm{Hom}_R(S_\kappa^{n.ord}(K,\varepsilon;R),R)$$
$$\text{and } S_\kappa^{n.ord}(K,\varepsilon;R) \cong \mathrm{Hom}_R(h_\kappa^{n.ord}(K,\varepsilon;R),R)$$
for a p-profinite algebra R finite over W;

(3) $$\mathbf{h}^{n.ord}(\mathfrak{N},\varepsilon;W[[T(\mathbb{Z}_p)]]) \cong \mathrm{Hom}_W(V_{cusp}^{ord}(\widehat{\Gamma}_0(\mathfrak{N}),\varepsilon;W),W)$$
$$\text{and } V_{cusp}^{ord}(\widehat{\Gamma}_0(\mathfrak{N}),\varepsilon;W) \cong \mathrm{Hom}_W(\mathbf{h}^{n.ord}(\mathfrak{N},\varepsilon;W[[T(\mathbb{Z}_p)]]),W);$$

(4) $$\mathbf{h}^{n.ord}(\mathfrak{N},\varepsilon;W[[T(\mathbb{Z}_p)]]) \cong \mathrm{Hom}_R(\mathcal{V}_{cusp}^{ord,*}(\widehat{\Gamma}_0(\mathfrak{N}),\varepsilon;W),R)$$
$$\text{and } \mathcal{V}_{cusp}^{ord,*}(\widehat{\Gamma}_0(\mathfrak{N}),\varepsilon;W) \cong \mathrm{Hom}_R(\mathbf{h}^{n.ord}(\mathfrak{N},\varepsilon;W[[T(\mathbb{Z}_p)]]),R)$$
for $R = W[[\Gamma]]$.

Proof. We give a sketch of a proof. The Hecke algebras and the spaces of modular forms involved in (1) and (2) are R-modules of finite rank. This follows from Proposition 4.27 for (1) and Theorem 4.9 for (2).

Since the proof of the assertions (1) and (2) is the same and (4) follows from (3), we prove only (1) and (3). By the long exact sequence of $\underline{\omega}_{\kappa,\varepsilon/W} \xrightarrow{p^m} \underline{\omega}_{\kappa,\varepsilon/W} \twoheadrightarrow \underline{\omega}_{\kappa,\varepsilon/W_m}$, $S_\kappa(K,\varepsilon;W)\otimes W_m$ $(= H^0(Sh_{\overline{K}/W},\underline{\omega}_{\kappa,\varepsilon/W})\otimes_W W_m)$ injects into $S_\kappa(K,\varepsilon;W_m)$ $(= H^0(Sh_{\overline{K}/W_m},\underline{\omega}_{\kappa,\varepsilon/W_m}))$; so, the q-expansion principle is valid for the finite module $S_\kappa(K,\varepsilon;W)\otimes W_m$. For the pairing $\langle\ ,\ \rangle : (h_\kappa(K,\varepsilon;W)\otimes W_m)\times(S_\kappa(K,\varepsilon;W)\otimes_W W_m)\to W_m$, we have $(U(y_\mathfrak{N})T(1,y^{(\mathfrak{N})}),f) = \mathbf{a}_p(y;f)$ by (4.67), and hence, by the q-expansion principle, $f=0$ if $(h_\kappa(K,\varepsilon;W)\otimes_W W_m,f)=0$. By the perfectness of the duality of finite modules, we thus have a surjective morphism

$$h_\kappa(K,\varepsilon;W)\otimes_W W_m \twoheadrightarrow \mathrm{Hom}_{W_m}(S_\kappa(K,\varepsilon;W)\otimes_W W_m, W_m)$$

of Hecke modules. Taking the projective limit relative to m, we get a surjective homomorphism $h_\kappa(K,\varepsilon;W)\twoheadrightarrow \mathrm{Hom}_W(S_\kappa(K,\varepsilon;W),W)$. Since $h_\kappa(K,\varepsilon;W)$ acts faithfully on $S_\kappa(K,\varepsilon;W)$, we conclude the injectivity. This implies

$$h_\kappa(K,\varepsilon;R)\cong \mathrm{Hom}_R(S_\kappa(K,\varepsilon;R),R) \quad \text{for } R=W.$$

Since W is faithfully flat over $\mathcal{W}$ and $\mathcal{W}$ is a valuation ring, we get at the same time $S_\kappa(K,\varepsilon;\mathcal{W})\cong \mathrm{Hom}_W(h_\kappa(K,\varepsilon;\mathcal{W}),\mathcal{W})$. If R is flat over $\mathcal{W}$, we find $S_\kappa(K,\varepsilon;R)=S_\kappa(K,\varepsilon;\mathcal{W})\otimes_{\mathcal{W}} R$ by flat base change (e.g., [ALG] III.9.3). This implies $h_\kappa(K,\varepsilon;R)=h_\kappa(K,\varepsilon;\mathcal{W})\otimes_{\mathcal{W}} R$ by the above two identities.

Replacing $h_\kappa(K,\varepsilon;W)\otimes_W W_m$ and $S_\kappa(K,\varepsilon;W)\otimes_W W_m$ by $h_\kappa(K,\varepsilon;W_m)$ and $S_\kappa(K,\varepsilon;W_m)$ in the above argument, we get the desired identity for $R=W_m$. Since any W-module of finite type is a finite direct sum of W_m-flat modules for some m (including $m=\infty$), the above argument using flat base change works well to get the desired result for a W-module R of finite type. This shows the assertion (1).

To prove the assertion (3), we simply write $\mathbf{h}$ for $\mathbf{h}^{n.ord}(\mathfrak{N},\varepsilon;W[[T(\mathbb{Z}_p)]])$. We consider the pairing $\langle\ ,\ \rangle : \mathcal{V}^{ord}_{cusp}(\widehat{\Gamma}_0(\mathfrak{N}))\times\mathbf{h}\to K/W$ given by $(f,h)=\mathbf{a}_p(1,f|h)$. Then again we have $(U(y_\mathfrak{N})T(1,y^{(\mathfrak{N})}),f)=\mathbf{a}_p(y;f)$ by (4.67) applied to p-adic modular forms and, hence, by the q-expansion principle, $\langle f,\mathbf{h}\rangle=0$ implies $f=0$. By the perfectness of the Pontryagin duality, we thus have a surjective morphism: $\mathbf{h}\twoheadrightarrow \mathcal{V}^{ord,*}_{cusp}(\widehat{\Gamma}_0(\mathfrak{N}))$ of Hecke modules. Since $\mathbf{h}$ acts faithfully on $\mathcal{V}^{ord,*}_{cusp}(\widehat{\Gamma}_0(\mathfrak{N}))$, we conclude the injectivity. Since we have $\mathcal{V}^{ord,*}_{cusp}(\widehat{\Gamma}_0(\mathfrak{N}))\cong \mathrm{Hom}_W(V^{ord}_{cusp}(\widehat{\Gamma}_0(\mathfrak{N}),\varepsilon;W),W)$ (e.g., Section 3.3.1 particularly (3.4)), assertion (3) follows. Assertion (4) is a formal consequence of (3) and the R-freeness of $\mathcal{V}^{ord,*}_{cusp}$. □

The above theorem combined with Theorem 4.12 implies

Corollary 4.31 *Let $\mathfrak{N}$ be an O-ideal with $\mathfrak{N}\cap\mathbb{Z}=(N)$. Suppose that $p\geq 5$ is prime to $Nd(F)$. Fix a $\kappa\in X(T_G)$ and a Neben character $\varepsilon=(\varepsilon_1,\varepsilon_2,\varepsilon_+)$ with ε^- factoring through $(O/\mathfrak{N})^\times$. Then we have for $T=\mathrm{Res}_{O/\mathbb{Z}}\mathbb{G}_m$,*

(1) *$\mathbf{h}^{n.ord}(\mathfrak{N},\varepsilon;W[[T(\mathbb{Z}_p)]])$ is a $W[[T(\mathbb{Z}_p)]]$-projective module of finite type;*

(2) *If* $\kappa_2' - \kappa_1' \geq 2I$ *for* $\kappa' \in \mathbb{Z}[I]^2$ *with* $\kappa_1' + \kappa_2' = \kappa_1 + \kappa_2$, *we have a canonical isomorphism of* W*-algebras*

$$\mathbf{h}^{n.ord}(\mathfrak{N}, \varepsilon; W[[T(\mathbb{Z}_p)]]) \otimes_{W[[T(\mathbb{Z}_p)]], \kappa'} W \cong h_{\kappa'}^{n.ord}(\mathfrak{N}, \varepsilon; W)$$

sending $T_p(y)$ *and* $U_p(y)$ *in* $\mathbf{h}^{n.ord}(\mathfrak{N}, \varepsilon; W[[T(\mathbb{Z}_p)]])$ *to* $T_p(y)$ *and* $U_p(y)$ *in* $h_{\kappa'}(\mathfrak{N}, \varepsilon; W)$. *If we replace* $\mathfrak{N}$ *by* $p\mathfrak{N}$ *in the right-hand side of the above identity, it is valid under the milder condition* $\kappa_2' - \kappa_1' \geq I$.

This corollary gives us a finite flat fibration $Spec(\mathbf{h}) \to D = Spec(W[[T(\mathbb{Z}_p)]])$ ($\mathbf{h} = \mathbf{h}^{n.ord}(S_0(\mathfrak{N}), \varepsilon; W[[T(\mathbb{Z}_p)]])$) whose fiber at an algebraic regular weight κ is the (spectrum of the) nearly ordinary Hecke algebra of weight κ with a fixed central character ε_+. Here the p-adic rigid analytic space of D is just a disjoint union of g-dimensional p-adic open unit ball ($g = [F : \mathbb{Q}]$); so, we have found a nice family of algebras from modular forms.

4.2.12 p-Adic Universal Hecke Algebra

We generalize Corollary 4.31 allowing the deformation of the central character. This means that we can expand the base D of the fibration $Spec(\mathbf{h}) \to D$ so that we can deform (or continuously vary) central characters of Hecke algebras appearing as its fiber. Let $Z = Cl_F^+(p^\infty)$ be the ray class group modulo $p^\infty\infty$, that is, $\varprojlim_r Cl_F^+(p^r)$. We decompose $Z = \Gamma_Z \times \Delta_Z$ so that Γ_Z is p-profinite and Δ_Z has order prime to p. Note that any character χ of Γ_Z can be regarded as a character of Z, because Z/Δ_Z is canonically isomorphic to Γ_Z. If $p > 2$, any character $\chi : \Gamma_Z \to W^\times$ has a unique square root $\sqrt{\chi}$. Thus we have a canonical identity $\underline{\omega}_{\kappa,\varepsilon} \otimes \sqrt{\chi} = \underline{\omega}_{\kappa,\varepsilon} \otimes_{W[[\Gamma_Z]], \sqrt{\chi}} W \cong \underline{\omega}_{\kappa, \varepsilon\otimes\chi}$ for $\varepsilon \otimes \chi = (\varepsilon_1\sqrt{\chi}, \varepsilon_2\sqrt{\chi}, \varepsilon_+\chi)$ (see (4.49)). We define

$$\mathbf{h}^{n.ord}(\mathfrak{N}, \varepsilon; W[[T(\mathbb{Z}_p) \times \Gamma_Z]]) = \mathbf{h}^{n.ord}(\mathfrak{N}, \varepsilon; W[[T(\mathbb{Z}_p)]]) \widehat{\otimes}_W W[[\Gamma_Z]], \tag{4.69}$$

which is called the *universal nearly p-ordinary Hecke algebra*. In the class of characters $\{\varepsilon_+\chi\}_\chi$ we have the Teichmüller lift ε_{0+} of ε_+ mod $\mathfrak{m}_W$. By construction, the above algebra only depends on the Teichmüller lift ε_0 of ε mod $\mathfrak{m}_W$ of weight $\kappa_0 = (0, I)$; so, we write the above algebra as $\mathbf{h}^{n.ord}(\mathfrak{N}, \varepsilon_0; W[[T(\mathbb{Z}_p) \times \Gamma_Z]])$.

For an idele y, we have a well-defined ideal class $\overline{y} \in Cl_F^+(p^\infty) = F_{\mathbb{A}}^\times / F^\times (\widehat{O}^{(p)})^\times F_{\infty+}^\times$. We write $[y]$ for the image of $\overline{y}$ under $Z \twoheadrightarrow \Gamma_Z$. We have the following consequence of Corollary 4.31 and the discussion in Section 4.2.7.

Corollary 4.32 *Suppose that* $p \nmid 2Nd(F)$ *and that* $p \geq 5$. *Then we have*

(1) *The Hecke algebra* $\mathbf{h}^{n.ord}(\mathfrak{N}, \varepsilon_0; W[[T(\mathbb{Z}_p) \times \Gamma_Z]])$ *is a* $W[[T(\mathbb{Z}_p) \times \Gamma_Z]]$*-projective module of finite type;*

(2) *When* $\kappa_2 - \kappa_1 \geq 2I$ *for* $\kappa \in X(T_G)$ *and* $\varepsilon|_{T_G(\widehat{\mathbb{Z}})} = \varepsilon_0|_{T_G(\widehat{\mathbb{Z}})}$ *with* $\varepsilon_+ \equiv \varepsilon_{0+}$ mod $\mathfrak{m}_W$ *on* $Z(\mathbb{A}^{(p\infty)})$, *for every arithmetic Hecke character* χ *of* Γ_Z *with* ∞*-type* $2v$, *we have a canonical isomorphism*

$$\mathbf{h}^{n.ord}(\mathfrak{N}, \varepsilon_0; W[[T(\mathbb{Z}_p) \times \Gamma_Z]]) \otimes_{W[[T(\mathbb{Z}_p)\times\Gamma_Z]],\kappa\otimes\sqrt{\chi}} W \cong h^{n.ord}_{\kappa+(v,v)}(\mathfrak{N}, \varepsilon \otimes \chi; W),$$

which sends $T_p(y)\otimes[y]$ *and* $U_p(y)\otimes[y]$ *in* $\mathbf{h}^{n.ord}(\mathfrak{N}, \varepsilon; W[[T(\mathbb{Z}_p) \times \Gamma_Z]])$ *to* $T_p(y)$ *and* $U_p(y)$ *in* $h^{n.ord}_{\kappa+(v,v)}(\mathfrak{N}, \varepsilon \otimes \chi; W)$. *If we replace* $\mathfrak{N}$ *by* $p\mathfrak{N}$ *in the right-hand side of the above identity, it is valid under the milder condition* $\kappa_2 - \kappa_1 \geq I$.

4.3 Rank of *p*-Ordinary Cohomology Groups

We give a sketch of a proof of the constancy of the rank of nearly ordinary cohomology groups when we vary the central character and the weight. The idea is the same as in the proof of Theorem 4.9 after moving everything to quaternionic automorphic forms by the Jacquet–Langlands and Shimizu correspondence. It is a well-known result of Jacquet–Langlands and Shimizu that if we choose the level appropriately, the space of holomorphic quaternionic automorphic forms can be embedded into the space of Hilbert modular forms preserving the Hecke operator action. If the ramification of the quaternion algebra is limited to Archimedean places, the embedding is surjective. Furthermore, if the Archimedean ramification is maximal, the Shimura variety has small dimension ≤ 1 so that we can make the argument more precise.

We first redefine the space of modular forms in 4.3.1 in order to state the correspondence in Section 4.3.2. In Section 4.3.3, we interpret the correspondence in terms of topological cohomology and make it canonical. Then we shall give in Section 4.3.4 an exposition of the Eichler–Shimura isomorphism which connects the topological cohomology groups and the space of cusp forms defined in Sections 4.3.1 and 4.3.2. At the end of this section, we give a proof of the boundedness in Section 4.3.5. *Contrary to* (ex0), *in this section* $y^w = \prod_{\sigma\in I}\sigma(y)^{w_\sigma}$ *for all* $y \in F_p^\times$ *and* $w \in \mathbb{Z}[I]$.

4.3.1 Archimedean Automorphic Forms

Let us recall the definition of the adelic Hilbert modular forms and their Hecke ring of level $\mathfrak{N}$ for an integral ideal $\mathfrak{N}$ of F (cf. [SGL] Sections 2.2–4). Let $T_0 = \mathbb{G}^2_{m/O}$ be the diagonal torus of $GL(2)_{/O}$; so, $T_G = \mathrm{Res}_{O/\mathbb{Z}}T_0$. Recall the characters $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_+)$ as in Section 4.2.5. In particular, $\varepsilon_+ : Z(\mathbb{A})/Z(\mathbb{Q}) \to \mathbb{C}^\times$ is an arithmetic Hecke character such that $\varepsilon_+(z) = \varepsilon_1(z)\varepsilon_2(z)$ and $\varepsilon_+(x_\infty) = x^{-(\kappa_1+\kappa_2)+I}$. Suppose that the conductor of ε_j $(j = 1, 2)$ is a factor of an integral ideal $\mathfrak{N}'$ and that the conductor of $\varepsilon^- = \varepsilon_2^{-1}\varepsilon_1$ is a factor

of $\mathfrak{N}|\mathfrak{N}'$. Recall the "central" character $\varepsilon_S^\infty : Z(\mathbb{A})S_0(\mathfrak{N}) \to \mathbb{C}^\times$ as in (ex3) given by $\varepsilon_S^\infty(zu) = \varepsilon_+(z)|z|_{\mathbb{A}}^{-2}\varepsilon_\Delta^\infty(u)$ for $u \in S_0(\mathfrak{N})$ and $z \in Z(\mathbb{A})$. To simplify the notation, we write ε for the restriction of ε_S^∞ to $S_0(\mathfrak{N})$. We have extended ε to $\Delta_0(\mathfrak{N})$ by $\varepsilon_\Delta^\infty(u) = \varepsilon_2(\det(u))\varepsilon^-(a_{\mathfrak{N}})$ for $u = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Delta_0(\mathfrak{N})$. This Neben type has been used in [F1].

For Archimedean automorphic forms $f : G(\mathbb{Q})\backslash G(\mathbb{A})/K \to \mathbb{C}$ of level K, it is standard to define Hecke operators by the sum $f|[KgK](x) = \sum_h f(gh)$ if $KgK = \bigsqcup_h hK$ without the determinant factor $|\det(g)|_{\mathbb{A}}$ (in (4.44)) in front of the sum $\sum_h$. To adjust our notation to this convention, we take the twist $(|\cdot|_{\mathbb{A}} \otimes f)(g) = |\det(g)|_{\mathbb{A}} f(g)$ as the object corresponding to a section $H^0(Sh_K(\mathbb{C}), \underline{\omega}_{\kappa,\varepsilon})$ if f restricted to $Sh_K(\mathbb{C})$ gives rise to the global section of $\underline{\omega}_{\kappa,\varepsilon}$. Indeed, we have $(|\cdot|_{\mathbb{A}} \otimes f)|[KgK](x) = |\cdot|_{\mathbb{A}} \otimes (|\det(g)|_{\mathbb{A}} \sum_h f(xh))$. This adjustment removes the factor $|\cdot|_{\mathbb{A}}^{-2}$ from the central character $|\cdot|_{\mathbb{A}}^{-2}\varepsilon_+$ of the sections of $\underline{\omega}_{\kappa,\varepsilon}$. Incorporating this adjustment, the automorphic factor $J_\kappa(g,z)$ of the weight κ in the sense of Section 1.1 satisfies $J_\kappa\left(\left(\begin{smallmatrix} t_1 & 0 \\ 0 & t_2 \end{smallmatrix}\right), z\right) = t_1^{\kappa_1 - I} t_2^{\kappa_2}$ for $\left(\begin{smallmatrix} t_1 & 0 \\ 0 & t_2 \end{smallmatrix}\right) \in T_G$ and is given by

$$J_\kappa(g,z) = \det(g)^{\kappa_1 - I} j(g,z)^{\kappa_2 - \kappa_1 + I} \quad \text{for } g \in G(\mathbb{R}) \text{ and } z \in \mathfrak{H}^I. \tag{4.70}$$

Then we redefine $S_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C})$ to be the space of functions $f : G(\mathbb{A}) \to \mathbb{C}$ satisfying the following conditions (e.g., [SGL] Section 2.2).

(SA1) We have the following automorphy

$$f(\alpha x u z) = \varepsilon_+(z)\varepsilon_\Delta^\infty(u) f(x) J_\kappa(u_\infty, \mathbf{i})^{-1}$$

for all $\alpha \in G(\mathbb{Q})$, $z \in Z(\mathbb{A})$, and $u \in S_0(\mathfrak{N})C_{\mathbf{i}}$ for the stabilizer $C_{\mathbf{i}}$ in $G(\mathbb{R})^+$ of $\mathbf{i} = (\sqrt{-1}, \ldots, \sqrt{-1}) \in \mathfrak{Z} = \mathfrak{H}^I$, where $G(\mathbb{R})^+$ is the identity-connected component of $G(\mathbb{R})$;

(SA2) Choosing $u \in G(\mathbb{R})$ with $u(\mathbf{i}) = z$ for each $z \in \mathfrak{H}^I$, define a function $f_g : \mathfrak{Z} \to \mathbb{C}$ by $f_g(z) = f(gu_\infty) J_\kappa(u_\infty, \mathbf{i})$ for each $g \in G(\mathbb{A}^{(\infty)})$. Then f_g is a holomorphic function on $\mathfrak{Z}$ for all g;

(SA3) $f_g(z)$ is exponentially decreasing as $\mathrm{Im}(z) \to \infty$.

As we now show, the right translations of the functions f satisfying (SA1–3) generate the (left) automorphic representation $f \mapsto |\det(g)|_{\mathbb{A}}(f|g^{-1})$ for $g \in G(\mathbb{A}^{(\infty)})$ whose central character is ε_+ (we studied in (4.61) its inverted right automorphic representation). If we replace the words "exponentially decreasing" in (S3) by "slowly increasing" at most polynomial growth as $\mathrm{Im}(z) \to \infty$, we rediscover the definition of the space of modular forms $G_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C})$.

The function f_g in (SA2) satisfies the classical automorphy condition

$$f_g(\gamma(z)) = \varepsilon^{-1}(g^{-1}\gamma g) f_g(z) J_\kappa(\gamma, z) \quad \text{for all } \gamma \in \Gamma_{0,g}(\mathfrak{N}), \tag{4.71}$$

where $\Gamma_{0,g}(\mathfrak{N}) = g \cdot S_0(\mathfrak{N}) g^{-1} G(\mathbb{R})^+ \cap G(\mathbb{Q})$, and $G(\mathbb{R})^+$ is the subgroup of $G(\mathbb{R})$ made up of matrices with a totally positive determinant. Indeed, we have under the notation in (SA2),

$$f_g(\gamma(z)) = f(g(\gamma u)_\infty)J_\kappa((\gamma u)_\infty, \mathbf{i}) = f(\gamma\gamma^{(\infty)^{-1}} g \cdot u_\infty)J_\kappa((\gamma u)_\infty, \mathbf{i})$$
$$= \varepsilon^{-1}(g^{-1}\gamma g)f(gu_\infty)J_\kappa(\gamma, u_\infty(\mathbf{i}))J_\kappa(u_\infty, \mathbf{i}) = \varepsilon^{-1}(g^{-1}\gamma g)f_g(z)J_\kappa(\gamma, z).$$

The same argument applied to $\alpha \in G(\mathbb{Q})^+$ yields

$$f_g(\alpha(z))J_\kappa(\alpha, z)^{-1} = f_{\alpha^{(\infty)^{-1}}g}(z). \tag{4.72}$$

By (SA3) combined with (4.72), f_g is exponentially decreasing towards all cusps of $\Gamma_{0,g}(\mathfrak{N})$. It is easy to see that $G_\kappa = 0$ unless $\kappa_1 + \kappa_2 = [\kappa_1 + \kappa_2]I$ for $[\kappa_1 + \kappa_2] \in \mathbb{Z}$. We simply write $[\kappa]$ for $[\kappa_1 + \kappa_2] \in \mathbb{Z}$ if $G_\kappa \neq 0$.

Let us now show that this new definition of $S_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C})$ is compatible with the earlier one, recalling the theory expounded in Section 4.1.3. We take $f'(A, \lambda, C, \omega)\omega^{\otimes k} \in H^0(Sh_{\overline{S}_0(\mathfrak{N})}, \underline{\omega}_{\kappa,\varepsilon/\mathbb{C}})$. The category of test objects (A, λ, C, ω) with polarization ideal $\mathfrak{c}$ is equivalent over $\mathbb{C}$ to the category of triples $(\mathcal{L}, \lambda : \mathcal{L} \wedge \mathcal{L} \cong \mathfrak{c}^*, C)$ for O-lattices $\mathcal{L}$ of $F_\mathbb{C} = \mathbb{C}^I$ and a subgroup C of $F_\mathbb{C}/\mathcal{L}$ cyclic of order $\mathfrak{N}$. The abelian variety A has a uniformization $A(\mathbb{C}) \cong F_\mathbb{C}/\mathcal{L}$ of complex manifolds. The differential ω corresponds to $du = \sum_\sigma du_\sigma$ for the variable $u = (u_\sigma)_{\sigma \in I}$. Since $\mathcal{L}$ is projective of rank 2 over O, we find two fractional ideals $(O, \mathfrak{c}^{-1})$ of F and a base (w_1, w_2) of $\mathcal{L}$ so that $\mathcal{L} = O^* w_1 + \mathfrak{c}^{-1} w_2$ with $z = \frac{w_2}{w_1} \in \mathfrak{Z}$ and $O^*\mathfrak{N}^{-1}/O^* = C$. The alternating form λ is given by $\lambda((-a \cdot w_1 + b \cdot w_2), (-a' \cdot w_1 + b' \cdot w_2)) = a'b - ab'$; so, the positivity $\mathrm{Im}(z) \gg 0$ implies the positivity of the induced polarization $\lambda : A \to {}^tA$ ([ABV] Chapter I). Multiplication by $\alpha = (2\pi i)w_2^{-1}$ sends $F_\mathbb{C}/\mathcal{L}$ to $F_\mathbb{C}/\mathcal{L}_z^\mathfrak{c}$ for $\mathcal{L}_z^\mathfrak{c} = \mathcal{L}_z$ in Section 4.1.3. We have $(\alpha^*)du = \alpha\omega$. Thus putting $f'(\mathcal{L}, \lambda, C) = f'(A, \lambda, C, \omega)$, we find $f'(\mathcal{L}, \lambda, C) = \alpha^{-k} f'(\mathcal{L}_z, \lambda_z, C_z)$ for $C_z = \alpha C = (O^*\mathfrak{N}^{-1} + \mathfrak{c}^{-1}z)/\mathcal{L}_z$. In other words, the restriction of the untwisted $|\cdot|_\mathbb{A}^{-1} \otimes f$ for an automorphic form f satisfying (SA1-3) to $\mathfrak{M}(\mathfrak{c}, \Gamma_0(\mathfrak{N}))(\mathbb{C})$ is determined by the function $f'(z) = f'(\mathcal{L}_z, \lambda_z, C_z)$ satisfying (4.71) because of (4.2) and (4.9) (combined with the action of $S_0(\mathfrak{N})$ given by ε). Taking $\widetilde{c} = \left(\begin{smallmatrix} c & 0 \\ 0 & 1 \end{smallmatrix}\right) \in G(\mathbb{A}^{(\infty)})$ with $c\widehat{O} \cap F = \mathfrak{c}$, we have $\Gamma_{0,\widetilde{c}}(\mathfrak{N}) = \Gamma_0(\mathfrak{c}, \mathfrak{N})$, and we find the tuple $\{f'_\mathfrak{c}(z)\}_\mathfrak{c}$ giving $\{|c|_\mathbb{A}^{-1} f_{\widetilde{c}}(z)\}_c$ for the corresponding Hilbert modular form $f : G(\mathbb{A}) \to \mathbb{C}$ as in (SA1–3).

By the cuspidality condition, f' is exponentially decreasing towards each cusp, since the q-expansion matches with the Fourier expansion substituting $\exp(\mathrm{Tr}(\xi z))$ for q^ξ. Thus we find a canonical embedding of

$$H^0(Sh_{\overline{S}_0(\mathfrak{N})}, \underline{\omega}_{\kappa,\varepsilon/\mathbb{C}}) = \bigoplus_\mathfrak{c} S_\kappa(\mathfrak{c}, \Gamma_0(\mathfrak{N}); \mathbb{C})$$

(as $\mathfrak{c}$ runs over $Cl^{[2]} = Cl_F^+/(Cl_F^+)^2$) into $S_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C})$ as in (SA1–3). A function f satisfying (SA1–3) gives a holomorphic section of the analytic line bundle $\underline{\omega}_{\kappa,\varepsilon}^{an}$ induced by $\underline{\omega}_{\kappa,\varepsilon/\mathbb{C}}$. Then by GAGA (see [Se] and [ALG] Appendix B), we have (from the quasi-projectivity of $Sh(G, X)(\mathbb{C})$)

$$S_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C}) = H^0(Sh_{\overline{S}_0(\mathfrak{N})}, \underline{\omega}_{\kappa,\varepsilon/\mathbb{C}}^{an}) \cong H^0(Sh_{\overline{S}_0(\mathfrak{N})}, \underline{\omega}_{\kappa,\varepsilon/\mathbb{C}}). \tag{4.73}$$

Thus the two definitions are compatible over $\mathbb{C}$.

For each $y \in \Delta_0(\mathfrak{N})$, we can decompose

$$S_0(\mathfrak{N})y^\iota S_0(\mathfrak{N}) = \bigsqcup_{u,t} utS_0(\mathfrak{N}) \tag{4.74}$$

for finitely many $u \in U(\widehat{\mathbb{Z}})$ and $t \in T_G(\mathbb{A}^{(\infty)})$ with $\det(t) = \det(y)$ (see [IAT] Chapter 3 or [MFG] 3.1.6). Since $\varepsilon_\Delta^\infty$ is well-defined over $\Delta_0(\mathfrak{N})$ and coincides with $z \mapsto \varepsilon_+(z_\mathfrak{N})\varepsilon_2(\det(z^{(\mathfrak{N})}))$ on $Z(\mathbb{A}^{(\infty)}) \cap \Delta_0(\mathfrak{N})$, we may extend it to the subgroup of $G(\mathbb{A}^{(\infty)})$ generated by $Z(\mathbb{A}^{(\infty)})\Delta_0(\mathfrak{N})$ (which contains $\Delta_0(\mathfrak{N})^\iota$) so that $\underline{\varepsilon}(z\delta) = \varepsilon_+(z_\mathfrak{N})\varepsilon_2(\det(z^{(\mathfrak{N})}))\varepsilon_\Delta^\infty(\delta)$ for $\delta \in \Delta_0(\mathfrak{N})$ and $z \in Z(\mathbb{A}^{(\infty)})$. We then have

$$f|[S_0(\mathfrak{N})y^\iota S_0(\mathfrak{N})](g) = \sum_{u,t} \varepsilon_\Delta^\infty((ut)^\iota) f(gut) \overset{(*)}{=} \underline{\varepsilon}(\det(y)) \sum_h \underline{\varepsilon}(ut)^{-1} f(gut). \tag{4.75}$$

The second identity $(*)$ follows from $\det(t) = \det(y)$, $(ut)^\iota = \det(t)(ut)^{-1}$ and multiplicativity of $\underline{\varepsilon}$, and by this, the sum is independent of the choice of u and t as long as $\det(t) = \det(y)$. It is easy to verify that the operator defined by (4.75) preserves the space $G_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C})$ and $S_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C})$ by confirming (SA1–3) for $f|[S_0(\mathfrak{N})y^\iota S_0(\mathfrak{N})]$.

The above description (4.75) of the Hecke operator $[S_0(\mathfrak{N})y^\iota S_0(\mathfrak{N})]$ looks different from the earlier definition in Sections 4.2.5 and 4.2.6 (especially (4.44)), because in the earlier definition, we used $x \mapsto x \cdot h^{-1}$ $(x \in Sh(G, X))$ instead of $g \mapsto g \cdot ut$ $(g \in G(\mathbb{A}))$. By applying the involution ι, we can take $h = (ut)^\iota$. We claim that the two definitions are the same (for the choice $v = \infty$ in (4.44)). Since the central characters are equal, the two actions coincide for each central element $z \in Z(\mathbb{A}^{(\infty)})$. Thus in order to show that the two operators coincide for all $[S_0(\mathfrak{N})y^\iota \cdot S_0(\mathfrak{N})]$ with $y \in \Delta_0(\mathfrak{N})$, we may assume that $y = \left(\begin{smallmatrix} 1 & 0 \\ 0 & \varpi_\mathfrak{q} \end{smallmatrix}\right)$. For such y, we have $t = y^\iota$ or $t = y$ in (4.74). If $t = y^\iota$, u runs over a complete representative set for $U(O/\mathfrak{q})$ in $U(O_\mathfrak{q})$, and if $t = y$, we may take $u = 1$. We take the decomposition

$$Sh_{S_0(\mathfrak{N})}(\mathbb{C}) = \bigsqcup_c G(\mathbb{Q})_+ \backslash \left(\mathfrak{Z} \times (G(\mathbb{Q})_+ \widetilde{c} S_0(\mathfrak{N}))\right) / S_0(\mathfrak{N}) \quad \left(\widetilde{c} = \left(\begin{smallmatrix} c & 0 \\ 0 & 1 \end{smallmatrix}\right)\right).$$

As before, we write $[z, g]$ for a point of $Sh(\mathbb{C})$ represented by $(z, g) \in \mathfrak{Z} \times G(\mathbb{A}^{(\infty)})$. We have $[z, \gamma g] = [\gamma^{-1}(z), g]$ for $\gamma \in G(\mathbb{Q})_+$. Thus taking $L_\mathfrak{c} = O^* \oplus \mathfrak{c}^{-1}$ and putting $\mathcal{L}_z = O^* + \mathfrak{c}^{-1} z$ for $z \in \mathfrak{Z}$, we have $\mathbb{X}_z(\mathbb{C}) = \mathbb{C}^I / \mathcal{L}_z$. By computation, if $\mathfrak{c}\varpi_\mathfrak{q} = \xi \mathfrak{c}'$ for $\mathfrak{c}'$ in the representative set for Cl_F^+, we find

$$[z, \widetilde{c} u y^\iota] = [z, \gamma_u \gamma_{y^\iota} \widetilde{c}' s_u] = [\gamma_{y^\iota}^{-1} \gamma_u^{-1}(z), \widetilde{c}' s_u]$$

for $s_u \in S_0(\mathfrak{N})$ and a finite idele c' with $c'\widehat{O} \cap F = \mathfrak{c}'$, where $\gamma_{y^\iota} = \left(\begin{smallmatrix} \xi & 0 \\ 0 & 1 \end{smallmatrix}\right)$ and $\gamma_u = \left(\begin{smallmatrix} 1 & v \\ 0 & 1 \end{smallmatrix}\right)$ with $v \in \mathfrak{c}$ running through the congruence classes in $\mathfrak{c}/\mathfrak{q}\mathfrak{c}$. Similarly we have

$$[z, \widetilde{c}y] = [\gamma_y^{-1}(z), \widetilde{c}'s]$$

with $s \in S_0(\mathfrak{N})$ and $\gamma_y \in G(\mathbb{Q})_+$. Since

$$\Gamma_0(\mathfrak{c}, \mathfrak{N})\gamma_y^{-1}\Gamma_0(\mathfrak{c}', \mathfrak{N}) = \Gamma_0(\mathfrak{c}, \mathfrak{N})\gamma_y^{-1} \sqcup \bigsqcup_u \Gamma_0(\mathfrak{c}, \mathfrak{N})\gamma_{y^\iota}^{-1}\gamma_u^{-1},$$

$\mathbb{X}_{\gamma_y^{-1}(z)}$ and $\mathbb{X}_{\gamma_{y^\iota}^{-1}\gamma_u^{-1}(z)}$ $(u \in U(O/\mathfrak{q}))$ give all the quotients of $\mathbb{X}_z$ by cyclic subgroups of order $\mathfrak{q}$. Indeed for $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(\mathfrak{c}, \mathfrak{N})\gamma_y^{-1}\Gamma_0(\mathfrak{c}', \mathfrak{N})$, we have a natural inclusion $\mathcal{L}_z \subset (cz + d)\mathcal{L}_{\gamma(z)} = L\gamma\left(\begin{smallmatrix} z \\ 1 \end{smallmatrix}\right)$ and an exact sequence

$$0 \to (cz + d)\mathcal{L}_{\gamma(z)}/\mathcal{L}_z \to \mathbb{X}_z \xrightarrow{(cz+d)^{-1}} \mathbb{X}_{\gamma(z)} \to 0$$

for $(cz + d)\mathcal{L}_{\gamma(z)}/\mathcal{L}_z \cong O/\mathfrak{q}$. Thus the two sets of abelian varieties,

$$\{\mathbb{X}_{\gamma_y^{-1}(z)}, \mathbb{X}_{\gamma_{y^\iota}^{-1}\gamma_u^{-1}(z)} \ (u \in U(O/\mathfrak{q}))\} \quad \text{and} \quad \{Y_h | S_0(\mathfrak{N})yS_0(\mathfrak{N}) = \sqcup_h S_0(\mathfrak{N})h\}$$

coincide, where Y_h is as in (4.44) made from $A = \mathbb{X}_z$. One can also check that this bijection extends to include the level structures, the polarizations and the differentials in (4.44) for $KgK' = S_0(\mathfrak{N})yS_0(\mathfrak{N})$. Thus the new action of $[S_0(\mathfrak{N})y^\iota S_0(\mathfrak{N})]$ for $y \in \Delta_0(\mathfrak{N})$ coincides with the one $[S_0(\mathfrak{N})yS_0(\mathfrak{N})]$ defined in (4.44). By linearity, this action of double cosets extends to the ring action of the double coset ring $R(S_0(\mathfrak{N}), \Delta_0(\mathfrak{N}))$.

Each member f of $S_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C})$ has a Fourier expansion of the following form ([MFG] Theorem 3.10 and [SGL] Sections 2.3–4),

$$f\left(\begin{smallmatrix} y & x \\ 0 & 1 \end{smallmatrix}\right) = |y|_{\mathbb{A}} \sum_{0 \ll \xi \in F} \mathbf{a}_\infty(\xi y, f)(\xi y_\infty)^{-\kappa_1}\mathbf{e}_F(i\xi y_\infty)\mathbf{e}_{\mathbb{A}}(\xi x). \tag{4.76}$$

Here $y \mapsto \mathbf{a}_\infty(y, f)$ is a function defined on $y \in F_{\mathbb{A}}^\times$ only depending on its finite part $y^{(\infty)}$, and $\left(\begin{smallmatrix} y_\infty & 0 \\ 0 & 1 \end{smallmatrix}\right) \mapsto y_\infty^{-\kappa_1}\mathbf{e}_F(iy_\infty)$ is the restriction of the canonical Whittaker function of $GL_2^+(\mathbb{R})$ to matrices of the form $\left(\begin{smallmatrix} * & 0 \\ 0 & 1 \end{smallmatrix}\right)$ (whose Mellin transform gives the optimal Γ-factor of the standard L-function of f). The function $\mathbf{a}_\infty(y, f)$ is supported by the set $(\widehat{O} \times F_\infty) \cap F_{\mathbb{A}}^\times$ of integral ideles. The determinant factor $|y|_{\mathbb{A}} = |\det\left(\begin{smallmatrix} y & x \\ 0 & 1 \end{smallmatrix}\right)|_{\mathbb{A}}$ in front of the summation $\sum_{0 \ll \xi \in F}$ comes from the twist we mentioned earlier.

In view of the decomposition

$$S_0(\mathfrak{N})\left(\begin{smallmatrix} \varpi_\mathfrak{q} & 0 \\ 0 & 1 \end{smallmatrix}\right)S_0(\mathfrak{N}) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & \varpi_\mathfrak{q} \end{smallmatrix}\right)S_0(\mathfrak{N}) \sqcup \bigsqcup_{u \in O_\mathfrak{q}/\mathfrak{q}} \left(\begin{smallmatrix} \varpi_\mathfrak{q} & u \\ 0 & 1 \end{smallmatrix}\right)S_0(\mathfrak{N})$$

for a prime $\mathfrak{q} \nmid \mathfrak{N}$, we can directly verify by computation (see [MFG] Corollary 3.11) the following Archimedean counterpart of (4.65),

$$\mathbf{a}_\infty(y, f|T(\varpi_\mathfrak{q})) = \begin{cases} \mathbf{a}_\infty(y\varpi_\mathfrak{q}, f) + N(\mathfrak{q})\varepsilon_+(\varpi_\mathfrak{q})\mathbf{a}_\infty(\frac{y}{\varpi_\mathfrak{q}}, f) & \text{if } \mathfrak{q} \nmid \mathfrak{N}, \\ \mathbf{a}_\infty(y\varpi_\mathfrak{q}, f) & \text{if } \mathfrak{q} | \mathfrak{N}. \end{cases} \tag{4.77}$$

We call an idele $y \in F_{\mathbb{A}}^{\times}$ *integral* if $y^{(\infty)} \in \widehat{O}$. For each $\mathbb{Q}$-algebra $R \subset \mathbb{C}$ containing the values of characters ε_J and κ on $T_G(\mathbb{Q})$, we have

$$S_\kappa(\mathfrak{N}, \varepsilon; R) = \left\{ f \in G_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C}) \middle| \mathbf{a}_\infty(y, f) \in R \text{ for all } y \right\}.$$

By definition, writing f_y for f_g in (S2) for $g = \begin{pmatrix} y^{(\infty)} & 0 \\ 0 & 1 \end{pmatrix}$, we see easily that

$$a(\xi, f_y) = \mathbf{a}_\infty(\xi y, f)\xi^{-\kappa_1}. \tag{4.78}$$

Then by the q-expansion principle, we find $S_\kappa(\mathfrak{N}, \varepsilon; R) = H^0(Sh_{\overline{S}_0(\mathfrak{N})}, \underline{\omega}_{\kappa,\varepsilon/R})$, and hence the definition is compatible with the earlier one. We restate (4.43) slightly more generally for a $\mathcal{W}$-algebra $R \subset \overline{\mathbb{Q}}$ with $i_\infty : \overline{\mathbb{Q}} \hookrightarrow \mathbb{C}$:

$$G_\kappa(\mathfrak{N}, \varepsilon; R) \otimes_{R, i_\infty} \mathbb{C} = G_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C}) \quad \text{and} \quad S_\kappa(\mathfrak{N}, \varepsilon; R) \otimes_{R, i_\infty} \mathbb{C} = S_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C}).$$

We recall the embedding $i_p : \overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_p$. Then for any $\overline{\mathbb{Q}}_p$-algebra R, we find, consistently with (4.43),

$$G_\kappa(\mathfrak{N}, \varepsilon; R) = G_\kappa(\mathfrak{N}, \varepsilon; \overline{\mathbb{Q}}) \otimes_{\overline{\mathbb{Q}}, i_p} R \text{ and } S_\kappa(\mathfrak{N}, \varepsilon; R) = S_\kappa(\mathfrak{N}, \varepsilon; \overline{\mathbb{Q}}) \otimes_{\overline{\mathbb{Q}}, i_p} R.$$

By linearity, $(y, f) \mapsto \mathbf{a}_\infty(y, f)$ extends to functions on $F_{\mathbb{A}}^{\times} \times G_\kappa(\mathfrak{N}, \varepsilon; R)$ with values in R. Then we rediscover from (4.78) the p-adic q-expansion coefficients $\mathbf{a}_p(y, f)$ of $f \in G_\kappa(\mathfrak{N}, \varepsilon; R)$ by the following formula,

$$\mathbf{a}_p(y, f) = y_p^{-\kappa_1} \mathbf{a}_\infty(y, f). \tag{4.79}$$

The formal q-expansion of f has values in the space of functions on $T(\mathbb{A}^{(\infty)})$ with values in the formal monoid algebra $R[[q^\xi]]_{\xi \in F_+^{\times}}$ of the multiplicative semi-group $F_+^{\times}$, which is given by

$$f(y) = \sum_{\xi \gg 0} \mathbf{a}_p(\xi y, f) q^\xi. \tag{4.80}$$

This is the p-adic analogue of the Archimedean Fourier expansion (4.76). In particular, if $\mathbf{a}_p(y, f) = 0$ for all integral ideles y, the modular form f vanishes. In other words, the q-expansion: $y \mapsto f(y)$ determines f uniquely (for any algebra R for which the space of R-integral modular forms is well-defined).

By the q-expansion principle, for any p-adically complete W-algebra R in $\widehat{\mathbb{Q}}_p$ (the p-adic completion of $\overline{\mathbb{Q}}_p$), we conclude

$$S_\kappa(\mathfrak{N}, \varepsilon; R) = \left\{ f \in S_\kappa(\mathfrak{N}, \varepsilon; \widehat{\mathbb{Q}}_p) \middle| \mathbf{a}_p(y, f) \in R \text{ for all integral } y \right\}.$$

As we have already seen in Section 4.2.5, these spaces have geometric meaning as the space of R-integral global sections of a line bundle of the Hilbert modular variety of level $\mathfrak{N}$, that is, $S_\kappa(\mathfrak{N}, \varepsilon; R) \cong H^0(Sh^{(p)}_{\overline{S}_0(\mathfrak{N})}, \underline{\omega}_{\kappa,\varepsilon/R})$ whose

right-hand side has meaning even for W-algebras not in $\widehat{\mathbb{Q}}_p$; so, we define $S_\kappa(\mathfrak{N}, \varepsilon; R)$ for any p-adically complete W-algebra by the right-hand side.

We hereafter assume

$$\text{either } p|\mathfrak{N} \text{ or } [\kappa] \geq 0, \tag{4.81}$$

in order to ensure the stability of the spaces of R-integral modular forms under Hecke operators (see Theorem 4.28).

Proposition 4.33 *Let R be an algebra for which $\kappa : T_G(R) \to R^\times$ and $\varepsilon : T_G(\widehat{\mathbb{Z}}) \to R^\times$ are well-defined. Suppose that the space of cusp forms $S_\kappa(\mathfrak{N}, \varepsilon; R)$ is well-defined. Write $H = h_\kappa(\mathfrak{N}, \varepsilon; R)$ and $S = S_\kappa(\mathfrak{N}, \varepsilon; R)$. Let V be an H-module and V' be an R-module of finite type with an R-bilinear product $\langle\ ,\ \rangle : V \times V' \to R$. Then we have*

(1) *The following formal q-expansion for $v \in V$ and $w \in V'$,*

$$f(v \otimes w)(y) = \sum_{\xi \gg 0} \langle v|T_p(\xi y), w\rangle q^\xi$$

gives a unique element of $S_\kappa(\mathfrak{N}, \varepsilon; R)$.

(2) *The map $v \otimes w \mapsto f(v \otimes w)$ gives an R-linear map of $V \otimes_R V'$ into S with $f((v|T_p(y)) \otimes w) = f(v \otimes w)|T_p(y)$. If furthermore, V' is an H-module and $\langle v|h, w\rangle = \langle v, w|h\rangle$ for all $v \in V$, $w \in V'$, and $h \in H$, then the map f induces an H-linear map $V \otimes_H V' \to S_\kappa(\mathfrak{N}, \varepsilon; R)$.*

(3) *Suppose that $\mathbb{T}$ is an R-algebra direct summand of H, and put $V(\mathbb{T}) = \mathbb{T}V$ and $S(\mathbb{T}) = \mathbb{T}S$. If $V(\mathbb{T})$ is $\mathbb{T}$-free of finite rank and $\mathrm{Hom}_R(V(\mathbb{T}), R)$ is embedded into V' by the pairing $\langle\ ,\ \rangle$, then the map $f : V(\mathbb{T}) \otimes_H V' \to S(\mathbb{T})$ is surjective.*

The formulation of this proposition was suggested by the expression of the theta correspondence given in [Sh8] II, Theorem 3.1.

Proof. We have an isomorphism $\iota : \mathrm{Hom}_R(H, R) \cong S$ given by $\mathbf{a}_p(y, \iota(\phi)) = \phi(T_p(y))$, which is an H-linear map (i.e., $\iota(\phi \circ h) = \iota(\phi)|h$; see the proof of Theorem 4.30). Since V is an H-module, $h \mapsto \langle v|h, w\rangle$ gives an element of $\mathrm{Hom}_R(H, R)$ and hence an element in S. The element has the expression as in (1) by the above explicit form of ι. The assertion (2) is then clear from (1). As for (3), by the isomorphism $\mathrm{Hom}_R(V(\mathbb{T}), R) \hookrightarrow V'$, each element of $\mathrm{Hom}(\mathbb{T}, R) \cong S(\mathbb{T})$ is a finite R-linear combination of $h \mapsto \langle v|h, w\rangle$ for $v \in V(\mathbb{T})$ and $w \in V'$; so, the surjectivity follows. □

4.3.2 Jacquet–Langlands–Shimizu Correspondence

In order to create an example of the module V in Proposition 4.33, we study cohomology groups on quaternionic Shimura varieties. See [62a], [63c], and [68c] in [CPS] and also [H94] and [H88] for more details of such cohomology.

Let D be a quaternion algebra over F, that is, a central simple algebra of dimension 4 over F (see [BNT] for generality of simple algebras and [AAQ] for the theory of quaternion algebras). We write G_D for the algebraic group defined over $\mathbb{Q}$ such that $G_D(R) = (D \otimes_{\mathbb{Q}} R)^\times$ for each $\mathbb{Q}$-algebra R. Thus if $D = M_2(F)$, we have $G = G_{M_2(F)}$. Let $d(D)^2$ be the discriminant of D; so, $d(D)$ is the product of primes $\mathfrak{q}$ such that $D_\mathfrak{q} = D \otimes_F F_\mathfrak{q}$ is a division algebra. We assume that $p \nmid d(D)$ and define subsets I_D and I^D of I by

$$D \otimes_{F,\sigma} \mathbb{R} \cong \begin{cases} M_2(\mathbb{R}) & \text{if } \sigma \in I_D \\ \mathbb{H} & \text{if } \sigma \in I - I_D = I^D, \end{cases} \tag{4.82}$$

where $\mathbb{H}$ is the Hamilton quaternion algebra over $\mathbb{R}$. Finite places dividing $d(D)$ and Archimedean places in I^D are called the ramified places for $D_{/F}$. We can create such a quaternion algebra as follows: Choose a quadratic extension K of F so that $K_\mathfrak{q} = K \otimes_F F_\mathfrak{q}$ is a field for each prime $\mathfrak{q}|d(D)$ and that $K_\sigma = K \otimes_{F,\sigma} \mathbb{R} = \mathbb{C}$ for $\sigma \in I^D$. Find $\delta \in F$ such that $\delta \in N_{K_\mathfrak{q}/F_\mathfrak{q}}(K_\mathfrak{q}^\times)$ if and only if the place $\mathfrak{q}$ is outside $d(D)$ and I^D. We find such a δ if and only if the number of ramified places of D is even (see [BNT] IX.3 and XIII.6), and $D \cong \left\{ \left(\begin{smallmatrix} a & b \\ -\delta b^\rho & a^\rho \end{smallmatrix}\right) \in M_2(K) | a, b \in K \right\}$, where ρ is the generator of $\mathrm{Gal}(K/F)$. By our choice of δ, we find $0 \neq x = \left(\begin{smallmatrix} a & b \\ -\delta b^\rho & a^\rho \end{smallmatrix}\right) \in D_\mathfrak{q}$ with $N(x) = aa^\rho + \delta bb^\rho = 0$ if and only if the place $\mathfrak{q}$ is outside I^D and $d(D)$. Thus $D_\mathfrak{q}$ is a division algebra if and only if $\mathfrak{q}$ divides $d(D)$ or in I^D. The map $N : D \to F$ and the trace of the matrix $\left(\begin{smallmatrix} a & b \\ -\delta b^\rho & a^\rho \end{smallmatrix}\right)$ are called the *reduced norm* and the *reduced trace* $\mathrm{Tr} : D \to F$ of D, respectively. By our construction, $D \hookrightarrow M_2(K)$ induces an isomorphism $D \otimes_F K \cong M_2(K)$ which sends $b \otimes k$ to kb.

We fix once and for all an extension of $\sigma : F \hookrightarrow \overline{\mathbb{Q}}$ to $\sigma : \overline{F} \cong \overline{\mathbb{Q}}$ for an algebraic closure $\overline{F}/F$. We take a quadratic extension K/F inside $\overline{F}$ so that $K \otimes_{F,\sigma} \mathbb{R} \cong \mathbb{R} \times \mathbb{R}$ as F-algebras for $\sigma \in I_D$, $K \otimes_F F_\mathfrak{p} \cong F_\mathfrak{p} \times F_\mathfrak{p}$ for primes $\mathfrak{p}|p$ and $D \otimes_F K \cong M_2(K)$. We can always choose such a quadratic extension K as long as $p \nmid d(D)$. These conditions imply $K \otimes_F \mathbb{R} \cong \mathbb{C}$ for $\sigma \in I^D$.

We identify $D \otimes_F K$ with $M_2(K)$ by the above isomorphism. We fix maximal O-orders O_D and O_K of D and K, respectively, and we suppose that $O_D \otimes_O O_K \subset M_2(O_K)$. Here an O-order $O_H \subset H$ in a finite-dimensional semi-simple F-algebra H is an O-subalgebra of H such that for every $x \in O_H$, the commutative subalgebra $O[x] \subset H$ is an O-module of finite type and $O_H \otimes_{\mathbb{Z}} \mathbb{Q} = H$. If O_H is an O-order, its conjugate $aO_H a^{-1}$ for $a \in H$ is another O-order. There are finitely many nonisomorphic maximal O-orders inside H up to conjugation.

For each prime ideal $\mathfrak{l}$ prime to $d(D)$, we fix an isomorphism $O_{D,\mathfrak{l}} \cong M_2(O_\mathfrak{l})$ so that for the p-adic place $\mathfrak{p}|p$ induced by $i_p \circ \sigma$, this isomorphism coincides with the one, $O_D \hookrightarrow M_2(O_K) \xrightarrow{i_p \circ \sigma} M_2(O_\mathfrak{p})$. By means of this isomorphism, we identify $D_\mathfrak{l}$ with $M_2(F_\mathfrak{l})$. For an integral ideal $\mathfrak{N}_0$ of F prime to $d(D)$, putting $\mathfrak{N} = \mathfrak{N}_0 d(D)$, we define

$$\widehat{\Gamma}_0^D(\mathfrak{N}) = \left\{ x \in \widehat{O}_D^\times | x_{\mathfrak{N}_0} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \text{ with } c \in \mathfrak{N}_0 O_{\mathfrak{N}_0} \right\}, \tag{4.83}$$

where $\widehat{O}_D = O_D \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}$, and $O_{\mathfrak{N}_0} = \prod_{\mathfrak{l}|\mathfrak{N}_0} O_{\mathfrak{l}}$. We set $S_0^D(\mathfrak{N}) = s^{-1}\widehat{\Gamma}_0^D(\mathfrak{N})s$ for $s = \left(\begin{smallmatrix} d^{(d(D))} & 0 \\ 0 & 1 \end{smallmatrix}\right)$ with a finite idele d such that $d\widehat{O} \cap F = \mathfrak{d}$. Similarly we define $\Delta_0^D(\mathfrak{N}) \subset D \otimes_{\mathbb{Q}} \mathbb{A}^{(\infty)}$ so that it is the product of local components $\Delta_{\mathfrak{l}}$ which coincide with the local components of $\Delta_0(\mathfrak{N})$ as long as $\mathfrak{l} \nmid d(D)$ and $\Delta_{\mathfrak{l}} = O_{D,\mathfrak{l}}$ if $\mathfrak{l}|d(D)$. Again we can think of the double coset ring $R(S_0^D(\mathfrak{N}), \Delta_0^D(\mathfrak{N})) \cong R(S_0^D(\mathfrak{N}), \Delta_0^D(\mathfrak{N})^\iota)$ ($x^\iota = N(x)x^{-1}$ for the reduced norm map $N : D \to F$). We have $\mathcal{T}(\mathfrak{l}) = S_0^D(\mathfrak{N})\left(\begin{smallmatrix} \varpi_{\mathfrak{l}} & 0 \\ 0 & 1 \end{smallmatrix}\right) S_0^D(\mathfrak{N})$ and $\mathcal{T}(\mathfrak{l},\mathfrak{l}) = S_0^D(\mathfrak{N})\varpi_{\mathfrak{l}} S_0^D(\mathfrak{N})$ in $R(S_0^D(\mathfrak{N}), \Delta_0^D(\mathfrak{N})^\iota)$ for $\mathfrak{l} \nmid d(D)$, because the local component $\Delta_0^D(\mathfrak{N})_{\mathfrak{l}}$ at $\mathfrak{l}$ of $\Delta_0^D(\mathfrak{N})$ is identical to $\Delta_0(\mathfrak{N})_{\mathfrak{l}}$. For $\mathfrak{l}|d(D)$, we take $\alpha_{\mathfrak{l}} \in O_{D,\mathfrak{l}}$ so that its reduced norm generates $\mathfrak{l}O_{\mathfrak{l}}$. Then we define $\mathcal{T}(\mathfrak{l}) = -S_0^D(\mathfrak{N})\alpha_{\mathfrak{l}} S_0^D(\mathfrak{N})$ for $\mathfrak{l}|d(D)$, and we have

$$R(S_0^D(\mathfrak{N}), \Delta_0^D(\mathfrak{N})) \cong R(S_0(\mathfrak{N}), \Delta_0(\mathfrak{N})). \tag{4.84}$$

The above isomorphism brings $\mathcal{T}(\mathfrak{l})$ and $\mathcal{T}(\mathfrak{l},\mathfrak{l})$ to the corresponding elements in the right-hand side. As an operator on the space of automorphic forms, $\mathcal{T}(\mathfrak{l},\mathfrak{l})$ induces the central action $\langle \mathfrak{l} \rangle$ of $\varpi_{\mathfrak{l}}$.

For a given ring R, we consider the following module $L(\kappa^*; R)$ of the multiplicative semi-group $M_2(R)$. Let $\kappa^* = (\kappa_1 + I, \kappa_2)$ and put $n = \kappa_2 - \kappa_1 - I \in \mathbb{Z}[I]$, which is the restriction of $\kappa^* \in X(T_G)$ to $T \subset G_1$, and we confirm $(\kappa|_T)^* = k - 2I = n$. The correspondence $\kappa \mapsto \kappa^*$ is the one for $G = \mathrm{Res}_{F/\mathbb{Q}} GL(2)$ we discussed in Section 1.1 (see Section 5.3.2 for more details of this fact). We suppose that $n \geq 0$ (i.e., $n_\sigma \geq 0$ for all $\sigma \in I$), and we consider polynomials with coefficients in R of $(X_\sigma, Y_\sigma)_{\sigma \in I}$ homogeneous of degree n_σ for each pair (X_σ, Y_σ). The collection of all such polynomials forms an R-free module $L(\kappa^*; R)$ of rank $\prod_\sigma (n_\sigma + 1)$.

As before, we write v for the fixed place p or ∞; so, the base ring $\mathcal{B}$ is W if $v = p$ and $\mathbb{C}$ if $v = \infty$. Suppose that R is a $\mathcal{B}$-algebra. Then $i_v(\sigma(\delta_v))$ (which we write simply $\sigma(\delta_v)$) for $\delta \in G_D(\mathbb{A})$ can be regarded as an element in $M_2(R)$. Take a Neben character ε as in (ex1–4) of Section 4.2.6 with $\varepsilon^-|_{T(\widehat{\mathbb{Z}})}$ factoring through $(O/\mathfrak{N}_0)^\times$. We define ε_j ($j = 1, 2$) as in (ex1) and extend ε to $\Delta_0^D(\mathfrak{N})$ by $\varepsilon_D^{\Delta}(\delta) = \varepsilon_2(N(\delta))\varepsilon^-(a)$ if $\delta_{\mathfrak{N}_0} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Since $s \mapsto \varepsilon_D^{\Delta}(s)$ and $z \mapsto \varepsilon_+(z)$ coincide on $Z(\mathbb{A}^{(\infty)}) \cap S_0^D(\mathfrak{N})$, we may extend ε to a character $\varepsilon_D^S : S_0^D(\mathfrak{N}) Z(\mathbb{A}^{(\infty)}) \to R^\times$ by $\varepsilon_D^S(zu) = \varepsilon_D^{\Delta}(u)\varepsilon_+(z)$ for $z \in Z(\mathbb{A}^{(\infty)})$ and $u \in S_0^D(\mathfrak{N})$. We let $\Delta_0^D(\mathfrak{N})$ and $Z(\mathbb{A}) S_0^D(\mathfrak{N}) G_D(\mathbb{R})^+$ act on $L(\kappa^*; R)$ as follows.

$$\begin{aligned} \delta \cdot \Phi\left(\left(\begin{smallmatrix} X_\sigma \\ Y_\sigma \end{smallmatrix}\right)\right) &= \varepsilon_D^{\Delta}(\delta)^{-1} N(\delta_v)^{\kappa_1} \Phi\left(\sigma({}^s(\delta_v)^\iota)\left(\begin{smallmatrix} X_\sigma \\ Y_\sigma \end{smallmatrix}\right)\right), \\ (zu) \cdot \Phi\left(\left(\begin{smallmatrix} X_\sigma \\ Y_\sigma \end{smallmatrix}\right)\right) &= \varepsilon_D^S(z^{(\infty)}u^{(\infty)})^{-1} N(z_v u_v)^{\kappa_1} \Phi\left(\sigma({}^s(z_v u_v)^\iota)\left(\begin{smallmatrix} X_\sigma \\ Y_\sigma \end{smallmatrix}\right)\right). \end{aligned} \tag{4.85}$$

Here $z \in Z(\mathbb{A})$, $u \in S_0^D(\mathfrak{N}) G_D(\mathbb{R})$, and $\delta^\iota = N(\delta)\delta^{-1}$ and ${}^s\delta = s\delta s^{-1}$ for s given just below (4.83). We write $L(\kappa^*\varepsilon; R)$ for the module $L(\kappa^*; R)$ with this $(\Delta_0^D(\mathfrak{N}), Z(\mathbb{A}) S_0^D(\mathfrak{N}) G_D(\mathbb{R}))$-action.

By computation, $z \in Z(\mathbb{A})$ acts on $L(\kappa^*\varepsilon; R)$ through scalar multiplication by $\varepsilon_+(z)^{-1} z_v^{\kappa_1 + \kappa_2 - I}$; in particular, $\epsilon \in O_+^\times \subset Z(\mathbb{A})$ acts trivially on $L(\kappa^*\varepsilon; R)$.

If S is a sufficiently small open compact subgroup so that $S \cap Z(\mathbb{Q}) \subset O_+^\times$ in $G(\mathbb{A}^{(\infty)})$, central elements in $\Gamma_x = xSx^{-1} \cap G_D(\mathbb{Q})$ act trivially on $L(\kappa^*\varepsilon; R)$.

Assume that $v = \infty$. We let $g \in G_D(\mathbb{R})$ with $N(g) \gg 0$ act on $\mathfrak{Z}_D = \mathfrak{H}^{I_D}$ by the linear fractional transformation of $g_\sigma = \sigma(g) \in GL_2(K \otimes_{K,\sigma} \mathbb{R}) = GL_2(\mathbb{R})$ component-wise. We put $C_{\sigma+}$ for the stabilizer of $\sqrt{-1}$ in the Lie group $(D \otimes_{F,\sigma} \mathbb{R})^\times$ and define a maximal compact subgroup $C_{\mathbf{i}}^D \subset G_D(\mathbb{R})$ modulo center by $\prod_{\sigma \in I_D} C_{\sigma+} \times \prod_{\sigma \in I^D} \mathbb{H}^\times$, which is the stabilizer in $G_D(\mathbb{R})$ of $\mathbf{i} = (\sqrt{-1}, \dots, \sqrt{-1}) \in \mathfrak{H}^{I_D}$. Thus we have $\mathfrak{Z}_D = \mathfrak{H}^{I_D} \cong G_D(\mathbb{R})^+/C_{\mathbf{i}}^D$ by $g(\mathbf{i}) \leftrightarrow g$ for the identity-connected component $G_D(\mathbb{R})^+$ of $G_D(\mathbb{R})$. Since G and G_D have the common center $Z = \mathrm{Res}_{F/\mathbb{Q}}\mathbb{G}_m$, we use the same symbol Z to indicate the center of G_D. For any open compact subgroup $S \subset S_0^D(\mathfrak{N})$, we think of the complex manifold associated with the Shimura variety:

$$Y(S) = Y^D(S) := G_D(\mathbb{Q})\backslash G_D(\mathbb{A})/Z(\mathbb{A})S \cdot C_{\mathbf{i}}^D.$$

We write simply $Y_0^D(\mathfrak{N})$ for $Y(S_0^D(\mathfrak{N}))$.

Since we have already defined the space of Hilbert modular forms over $\mathbb{C}$ in (SA1–3), assuming that D is a division algebra, we state a definition of automorphic forms on $G_D(\mathbb{A})$. For each κ, we define $\kappa^D \in \mathbb{Z}[I]^2$ by $\kappa^D = (\kappa_1^D, \kappa_2^D)$ for $\kappa_j^D = \sum_{\sigma \in I^D} \kappa_{j,\sigma}\sigma$. Thus κ^D is the projection of κ to $\mathbb{Z}[I^D]^2$. Similarly, we define κ_D by the projection of κ to $\mathbb{Z}[I_D]^2 \subset \mathbb{Z}[I]^2$.

With each $\kappa \in \mathbb{Z}[I]^2 = X(T_G)$, we associate an automorphic factor,

$$J_\kappa^D(g, z) = \det(g)^{\kappa_{D,1} - I_D} j(g, z)^{\kappa_{D,2} - \kappa_{D,1} + I_D}, \tag{4.86}$$

for $g \in G_D(\mathbb{R})$ and $z \in \mathfrak{Z}_D$. We write $\kappa^{*,D}$ for the projection of κ^* to $\mathbb{Z}[I^D]$. Then, if $I_D \neq \emptyset$, we define $S_\kappa^D(\mathfrak{N}, \varepsilon; \mathbb{C})$ to be the space of functions f on $G_D(\mathbb{A})$ with values in $L_{C_{\mathbf{i}}^D}(\kappa^{*,D}, \mathbb{C})$ satisfying the following conditions.

(SB1) We have the following automorphy

$$f(\alpha x u z) = \varepsilon_+(z)\varepsilon_D^{\Delta}(u^{(\infty)})u_\infty^{-1} \cdot f(x) J_\kappa^D(u_\infty, \mathbf{i})^{-1}$$

for all $\alpha \in G_D(\mathbb{Q})$, $z \in Z(\mathbb{A})$, and $u \in S_0^D(\mathfrak{N})C_{\mathbf{i}}$, where $G_D(\mathbb{R})^+$ is the identity-connected component of $G_D(\mathbb{R})$. Here $f(x) \mapsto u_\infty \cdot f(x)$ is the action of the I^D-component u_∞^D of u_∞ on $L_{C_{\mathbf{i}}^D}(\kappa^{*,D}; \mathbb{C})$;

(SB2) Choosing $u \in G_D(\mathbb{R})$ with $u^D = 1$ and $u(\mathbf{i}) = z$ for each $z \in \mathfrak{Z}_D$, define a function $f_g : \mathfrak{Z}_D \to \mathbb{C}$ by $f_g(z) = f(gu_\infty)J_\kappa(u_\infty, \mathbf{i})$ for each $g \in G(\mathbb{A}^{(\infty)})$. Then f_g is a holomorphic function on $\mathfrak{Z}_D$ for all g.

When $I^D = I$ ($\Leftrightarrow I_D = \emptyset$), the variety $Y^D(S)$ is a finite set of points; so, the condition (SB2) is empty. However, writing $\mathcal{S}_\kappa^D(\mathfrak{N}, \varepsilon; \mathbb{C})$ for the space of functions satisfying (SB1) in this definite case, we need to take $S_\kappa^D(\mathfrak{N}, \varepsilon; \mathbb{C})$ to be the following quotient: $S_\kappa^D(\mathfrak{N}, \varepsilon; \mathbb{C}) = \mathcal{S}_\kappa^D(\mathfrak{N}, \varepsilon; \mathbb{C})/Iv(\mathfrak{N}, \varepsilon; \mathbb{C})$, where $Iv(\mathfrak{N}, \varepsilon; \mathbb{C})$ is the subspace made up of functions in $\mathcal{S}_\kappa^D(\mathfrak{N}, \varepsilon; \mathbb{C})$ factoring through the reduced norm map $N : G_D(\mathbb{A}) \to T(\mathbb{A})$. If $\kappa_2 - \kappa_1 \neq I$ or ε is nontrivial for some $x \in G_D(\mathbb{A})$ with $N(x) = 1$, $Iv(\mathfrak{N}, \varepsilon; \mathbb{C}) = 0$; so, no

modification is necessary. Decomposing $S_0^D(\mathfrak{N})y^\iota S_0^D(\mathfrak{N})$ for $y \in \Delta_0^D(\mathfrak{N})$ into $\bigsqcup_h hS_0^D(\mathfrak{N})$, we shall define the action of $R(S_0^D(\mathfrak{N}), \Delta_0^D(\mathfrak{N}))$ on $S_\kappa^D(\mathfrak{N}, \varepsilon; \mathbb{C})$ by

$$f|[S_0^D(\mathfrak{N})y^\iota S_0^D(\mathfrak{N})](g) = \sum_h h \cdot f(gh) = \sum_h h^{-\iota} \cdot f(gh^{-\iota}). \tag{4.87}$$

We state the theorem of Jacquet–Langlands and Shimizu in the following way.

Theorem 4.34 *Suppose $\mathfrak{N} = \mathfrak{N}_0 d(D)$ for an integral ideal $\mathfrak{N}_0$ prime to $d(D)$. Identify $R(S_0(\mathfrak{N}), \Delta_0(\mathfrak{N}))$ and $R(S_0^D(\mathfrak{N}), \Delta_0^D(\mathfrak{N}))$ as in* (4.84). *Then we have an $R(S_0(\mathfrak{N}), \Delta_0(\mathfrak{N}))$-linear embedding $S_\kappa^D(\mathfrak{N}, \varepsilon; \mathbb{C}) \hookrightarrow S_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C})$ for all dominant regular κ with $\kappa_2^D - \kappa_1^D \geq I^D$. The image of this embedding only depends on $d(D)$ and is made up of cusp forms in $S_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C})$ new at all primes $\mathfrak{q}|d(D)$. In particular, if $d(D) = 1$, the above morphism is a surjective isomorphism.*

Since known proofs of this theorem require harmonic analysis and a good knowledge of representation theory of adele groups (especially the theta correspondence), the proof is beyond the scope of this rather algebraic book; so, we quote some references where one can find a proof. The above formulation of the theorem was given in [H81] Section 2 and [H88] Theorem 2.1, where one can find an exposition of how to deduce this result from the original result of Jacquet–Langlands (whose exposition can be found in [AAG] Section 10). We hope to give an exposition of this result in a forthcoming book (through a line closer to Shimizu's treatment).

4.3.3 Integral Correspondence

Since the isomorphism of Theorem 4.34 is noncanonical, we rigidify it using some ideas of Shimizu and Shimura. Although we can treat the general case in a uniform manner, introducing the cuspidal cohomology group, in order to avoid technicality, we assume at the end of this section that $|I_D| \leq 1$, which is the case we need later. For the moment, we do not impose any condition on D; also, v can be any of the places p and ∞. Suppose that S is sufficiently small so that the image $\overline{\Gamma}_{S,x}$ of $\Gamma_{S,x} = xSx^{-1}G_D^+(\mathbb{R}) \cap G_D(\mathbb{Q})$ in $PG_D(\mathbb{R})$ acts freely on $\mathfrak{H}^{I_D}$ for all $x \in G_D(\mathbb{A}^{(\infty)})$ and that the action of $\Gamma_{S,x}$ on $L(\kappa^*\varepsilon; R) := L_{G_D}(\kappa^*\varepsilon; R)$ factors through $\overline{\Gamma}_{S,x}$. As already remarked, the central action of $\epsilon \in Z(\mathbb{Q}) \cap S \cdot G_D(\mathbb{R})$ on $L(\kappa^*\varepsilon; R)$ is the multiplication by $\varepsilon_+(\epsilon^{(\infty)})^{-1}\epsilon^{\kappa_1+\kappa_2-I}$, which has to be trivial. Then we can define an étale space over $Y(S)$: $\mathcal{L}(\kappa\varepsilon; R) = G_D(\mathbb{Q})\backslash (G_D(\mathbb{A}) \times L(\kappa^*\varepsilon; R))/Z(\mathbb{A})S \cdot C_{\mathbf{i}}^D$, where $\gamma(x, \Phi)zu = (\gamma xzu, (zu)^{-1} \cdot \Phi))$ for $u \in S \cdot C_{\mathbf{i}}^D$, $z \in Z(\mathbb{A})$, and $\gamma \in G_D(\mathbb{Q})$. This étale space gives rise to a sheaf $L(\kappa^*\varepsilon; R)_{/Y(S)}$ of locally constant sections. We write $H^q(Y(S), L(\kappa^*\varepsilon; R))$ for the cohomology group of $Y(S)$ with coefficients in the locally constant sheaf $L(\kappa^*\varepsilon; R)$.

Since $Y(S) \cong \sqcup_x \overline{\Gamma}_x \backslash \mathfrak{H}^{I_D}$ for finitely many x with $x_v = 1$, we have a canonical isomorphism (cf. [LFE] Appendix and [H94] page 470),

$$H^q(Y(S), L(\kappa^*\varepsilon; R)) \cong \bigoplus_x H^q(\overline{\Gamma}_{S,x}, L(\kappa^*\varepsilon; R)), \tag{4.88}$$

where the right-hand side is the direct sum of the group cohomology of the $\overline{\Gamma}_x$-module $L(\kappa^*\varepsilon; R)$. Even if $\overline{\Gamma}_{S,x}$ does not act freely on the module $L(\kappa^*\varepsilon; R)$, we still have $Y(S) \cong \bigsqcup_x \overline{\Gamma}_x \backslash \mathfrak{H}^{I_D}$ for finitely many x with $x_p = 1$, and we *define* the left-hand side of (4.88) by the right-hand side of (4.88).

If $S \subset S'$, we have the trace map $\mathrm{Tr}_{S'/S}$ (i.e., the transfer map in group cohomology; see [MFG] 4.3.1) and the restriction map $\mathrm{Res}_{S'/S}$:

$$\begin{aligned}\mathrm{Tr}_{S'/S} &: H^q(Y^D(S), L(\kappa^*\varepsilon; R)) \to H^q(Y^D(S'), L(\kappa^*\varepsilon; R)),\\ \mathrm{Res}_{S'/S} &: H^q(Y^D(S'), L(\kappa^*\varepsilon; R)) \to H^q(Y^D(S), L(\kappa^*\varepsilon; R)).\end{aligned}$$

We choose S sufficiently small so that $[S_0^D(\mathfrak{N}) : S]$ is prime to p (this is a condition on p). Again we have Tr and Res:

$$\begin{aligned}\mathrm{Tr} &: H^q(Y^D(S), L(\kappa^*\varepsilon; R)) \to H^q(Y_0^D(\mathfrak{N}), L(\kappa^*\varepsilon; R)),\\ \mathrm{Res} &: H^q(Y_0^D(\mathfrak{N}), L(\kappa^*\varepsilon; R)) \to H^q(Y^D(S), L(\kappa^*\varepsilon; R)).\end{aligned}$$

Since $\mathrm{Tr} \circ \mathrm{Res}$ is the scalar multiplication by $[S_0(\mathfrak{N}) : S]$ (cf. [MFG] 4.21), we have

$$H^q(Y_0^D(\mathfrak{N}), L(\kappa^*\varepsilon; R)) = H^q(Y^D(S), L(\kappa^*\varepsilon; R))/\operatorname{Ker}(\mathrm{Tr}) = \operatorname{Im}(\mathrm{Res}).$$

We can always choose a multiple $\mathfrak{N}' = \mathfrak{N}\mathfrak{q}$ (by a prime $\mathfrak{q}$) of $\mathfrak{N}$ so that $\overline{\Gamma}_{0,x}(\mathfrak{N}')$ acts freely on $\mathfrak{H}^{I_D}$.

We have a natural action of the ring $R(S_0^D(\mathfrak{N}), \Delta_0^D(\mathfrak{N}))$ on the cohomology group $H^q(Y_0^D(\mathfrak{N}), L(\kappa^*\varepsilon; R))$. The definition of the action is quite similar to the one (4.37) given for the coherent cohomology groups $H^0(Sh_K, \underline{\omega}_\varepsilon^k)$, and the operator induced by the double coset SgS for $S = S_0^D(\mathfrak{N})$ is given by

$$[S\delta S]_v = N(\delta_v)^{-\kappa_1}(\mathrm{Tr}_{S/(\delta S \cap S)} \circ [\delta] \circ \mathrm{Res}_{S/(S \cap S^\delta)}). \tag{4.89}$$

Here $[\delta] : L(\kappa^*\varepsilon; R) \to L(\kappa^*\varepsilon; R)$ is induced from the action of δ on $L(\kappa^*\varepsilon; R)$. If $q = 0$, decomposing $S_0^D(\mathfrak{N}) \left(\begin{smallmatrix} y & 0 \\ 0 & 1 \end{smallmatrix}\right) S_0^D(\mathfrak{N}) = \bigsqcup_\varpi \varpi S_0^D(\mathfrak{N})$, we have

$$f|T_v(1, y) = y_v^{-\kappa_1} \sum_\varpi \varpi f(x\varpi). \tag{4.90}$$

Suppose $I_D = \emptyset$. Then $Y^D(S)$ is a finite set, and $C_{\mathfrak{i}} = G_D(\mathbb{R})$. For a W-algebra R, we may identify $H^0(Y^D(S), L(\kappa^*\varepsilon; R))$ with the space of functions $f : G(\mathbb{A}) \to L(\kappa^*; R)$ satisfying

(SB) We have the automorphy $f(\alpha x u z) = \varepsilon_+(z)\varepsilon^{\Delta_D}(u^{(\infty)})u_v^{-1} \cdot f(x)$ for all $\alpha \in G_D(\mathbb{Q})$, $z \in Z(\mathbb{A})$, and $u \in S_0^D(\mathfrak{N})C_{\mathfrak{i}}^D$.

Since f satisfying (SB) is left invariant under $G_D(\mathbb{R})$, we may regard it as a function on $G(\mathbb{A}^{(\infty)}) = G(\mathbb{A})/G(\mathbb{R})$. Recall $Iv(\kappa, \varepsilon; R)$ which is made up of

functions as in (SB) factoring through the reduced norm $N : G(\mathbb{A}) \to T(\mathbb{A})$. We have, if $I_D = \emptyset$,

$$S_\kappa(\mathfrak{N}, \varepsilon; R) = H^0(Y_0^D(\mathfrak{N}), L(\kappa^*\varepsilon; R))/Iv(\mathfrak{N}, \varepsilon; R). \tag{4.91}$$

If R is p-torsion (so, $\mathcal{B} = W$ and $v = p$), $Iv(\mathfrak{N}, \varepsilon; R)$ could be nontrivial even if $\kappa_2 - \kappa_1 \neq I$, because $\kappa_2 - \kappa_1 - I$ could be trivial as a character of $T(R)$ (even if it is nontrivial as a character of the algebraic group T).

We assume that $R = W$ and $v = p$. We have p^g right cosets ($g = [F : \mathbb{Q}]$) in $\left(S_0^D(p\mathfrak{N})\left(\begin{smallmatrix} p_p & 0 \\ 0 & 1 \end{smallmatrix}\right) S_0^D(p\mathfrak{N})\right)/S_0^D(p\mathfrak{N})$, and each coset acts on $f \in Iv(\mathfrak{N}, \varepsilon; R)$ by $f(x) \mapsto f(xp_p)$, regarding f as a function of $T(\mathbb{A}^{(\infty)})$. Since f actually factors through a finite quotient of $T(\mathbb{A}^{(\infty)})$, for a high power $U(p)^n = U(p^n)$ of $U_p(p)$, we have $f|U(p^n) = p^{ng}f$; so, $Iv(\mathfrak{N}, \varepsilon; R)$ is killed by the p-ordinary projector e. Thus we have

$$S_\kappa^{n.ord}(p\mathfrak{N}, \varepsilon; R) = H^0_{n.ord}(Y_0^D(p\mathfrak{N}), L(\kappa^*\varepsilon; R)), \tag{4.92}$$

where the superscript or the subscript $n.ord$ indicates the image of the projector $e = \lim_{n\to\infty} U(p)^{n!}$. Since $T(p) \equiv U(p) \mod p$ if $\kappa_2 - \kappa_1 \geq I$ and $[\kappa] \geq 1$, the same assertion holds for $\mathfrak{N}$ prime to p:

$$S_\kappa^{n.ord}(\mathfrak{N}, \varepsilon; R) = H^0_{n.ord}(Y_0^D(\mathfrak{N}), L(\kappa^*\varepsilon; R)) \text{ if } \kappa_2 - \kappa_1 \geq I \text{ and } [\kappa] \geq 1.$$

Hereafter we assume that $|I_D| \leq 1$ and $F \neq \mathbb{Q}$. Write $q = |I_D|$. In this case, D is a division algebra, and $Y(S)$ is a compact manifold (cf. [Bo] 1.5). Let V be the image of $H^q(Y_0^D(\mathfrak{N}), L(\kappa^*\varepsilon; W))$ in $H^q(Y_0^D(\mathfrak{N}), L(\kappa^*\varepsilon; W \otimes_{\mathbb{Z}} \mathbb{Q}))$. By the Eichler–Shimura isomorphism (described in the following subsection) between the space of cusp forms on $G_D(\mathbb{A})$ and the cohomology group (e.g., [H94] Proposition 3.1 and (10.4)) combined with Theorem 4.34, the above cohomology group V is a module over the Hecke algebra $h_\kappa(\mathfrak{N}, \varepsilon; W)$.

We consider the duality pairing $[\ ,\ ]$ on $L(\kappa^*\varepsilon; R)$ (for $\mathbb{Q}$-algebra R) introduced in [H94] Section 5,

$$\left[\sum_{0\leq \jmath\leq n} b_\jmath X^{n-\jmath}Y^\jmath, \sum_{0\leq \jmath\leq n} a_\jmath X^{n-\jmath}Y^\jmath\right] = \sum_j (-1)^j \binom{n}{j}^{-1} b_{n-j}a_j \in R, \tag{4.93}$$

where $n = \kappa_2 - \kappa_1 - I$, $\binom{n}{\jmath} = \prod_{\sigma\in I}\binom{n_\sigma}{\jmath_\sigma}$, and, for example, $X^\jmath = \prod_{\sigma\in I} X_\sigma^{\jmath_\sigma}$. As $S_0(\mathfrak{N})Z(\mathbb{A})$-modules, this pairing satisfies:

$$[uz\Phi, uz\Phi'] = \varepsilon_D^{\Delta}(u^{(\infty)})^2\widehat{\varepsilon}_+(z)^2 N(u_p)^{([\kappa]-1)I}[\Phi, \Phi'], \tag{4.94}$$

where $N : D_p \to F_p$ is the reduced norm map and $\widehat{\varepsilon}_+(z) = \varepsilon_+(z)z_p^{-\kappa_1-\kappa_2+I}$ (the p-adic avatar of ε_+).

Define $\widehat{\kappa} = (I - \kappa_2, I - \kappa_1)$. Then the pairing $[\ ,\ ]$ induces $S_0(\mathfrak{N})Z(\mathbb{A})$-equivariant pairing $[\ ,\ ] : L(\kappa^*\varepsilon; R) \times L(\widehat{\kappa}\varepsilon^{-1}; R) \to R$. Recall $q = |I_D| =$

$\dim_{\mathbb{C}} \mathfrak{H}^{I_D}$. Then by the compactness of $Y_0^D(\mathfrak{N})$, the cup product pairing induces ([H94] (5.3)), a nondegenerate pairing

$$(\ , \ ) : H^q(Y_0^D(\mathfrak{N}), L(\kappa^*\varepsilon; W\otimes_{\mathbb{Z}}\mathbb{Q}))\times H^q(Y_0^D(\mathfrak{N}), L(\widehat{\kappa}\varepsilon^{-1}; W\otimes_{\mathbb{Z}}\mathbb{Q})) \to W\otimes_{\mathbb{Z}}\mathbb{Q}.$$

Thus we obtain from Proposition 4.33 and Theorem 4.36 the following result.

Proposition 4.35 *Let $V \subset H^q(Y_0^D(\mathfrak{N}), L(\kappa^*\varepsilon; W \otimes_{\mathbb{Z}} \mathbb{Q}))$ be as above for $q = |I_D| \leq 1$. Suppose $F \neq \mathbb{Q}$. Let V^* be the dual W-lattice of V in the cohomology group $H^q(Y_0^D(\mathfrak{N}), L(\widehat{\kappa}\varepsilon^{-1}; W \otimes_{\mathbb{Z}} \mathbb{Q}))$ under the Poincaré duality pairing $(\ , \ )$ as above. Then we have an $h_\kappa(\mathfrak{N}, \varepsilon; W)$-linear map $f : V\otimes_W V^* \to S_\kappa(\mathfrak{N}, \varepsilon; W)$ defined by the q-expansion: $f(v \otimes w) = \sum_{\xi \gg 0}(v|T_p(\xi y), w)q^\xi$, where we regard the space $V \otimes_W V^*$ as an $h_\kappa(\mathfrak{N}, \varepsilon; W)$-module through the left factor V.*

A similar fact for the matrix coefficients of $T(y)$ in place of $(v|T(y), w)$ has been proven in [Sh8] II, Theorem 3.1 by analytic means without using the Jacquet–Langlands correspondence.

4.3.4 Eichler–Shimura Isomorphisms

In this subsection, we assume $\mathcal{B} = \mathbb{C}$ and $v = \infty$. We relate $S_\kappa^D(\mathfrak{N}, \varepsilon; \mathbb{C})$ and $H^q(Y_0^D(\mathfrak{N}), L(\kappa^*\varepsilon; \mathbb{C}))$. The isomorphism is called the Eichler–Shimura isomorphism. When $|I_D| = 0$, (4.91) gives a direct relation; so, there is nothing left to do. Thus we assume $|I_D| = 1$ and $F \neq \mathbb{Q}$. Then $Y_0^D(\mathfrak{N})$ is a Riemann surface, which is a disjoint union of $\Gamma_{0,x}(\mathfrak{N})\backslash\mathfrak{H}$ for a suitable finite set of $x \in G_D(\mathbb{A}^{(\infty)})$. Thus Shimura's theory expounded in [IAT] Chapter 8 applies, which yields the following canonical isomorphism,

$$\begin{aligned} S_\kappa(\Gamma_{0,x}(\mathfrak{N}), \varepsilon; \mathbb{C}) \oplus \overline{S}_\kappa(\Gamma_{0,x}(\mathfrak{N}), \varepsilon; \mathbb{C}) &\cong H^0(\Gamma_{0,x}(\mathfrak{N}), L(\kappa^*\varepsilon; \mathbb{C})) \\ &\cong H^0(\Gamma_{0,x}(\mathfrak{N})\backslash\mathfrak{H}, L(\kappa^*\varepsilon; \mathbb{C})), \end{aligned} \tag{4.95}$$

where $\overline{S}_\kappa(\Gamma_{0,x}(\mathfrak{N}), \overline{\varepsilon}; \mathbb{C})$ is made up of complex conjugates of functions in the space $S_\kappa(\Gamma_{0,x}(\mathfrak{N}), \overline{\varepsilon}; \mathbb{C})$, and $S_\kappa(\Gamma_{0,x}(\mathfrak{N}), \varepsilon; \mathbb{C})$ is the collection of holomorphic functions $f : \mathfrak{H} \to L_C(\kappa^{*,D}, \mathbb{C})$ satisfying the following automorphy condition: $f(\gamma(z)) = \varepsilon^{-1}(x^{-1}\gamma x)\gamma \cdot f(z)J_\kappa^D(\gamma, z)$ for all $\gamma \in \Gamma_{0,x}(\mathfrak{N})$. The fact $S_\kappa^D(\mathfrak{N}, \varepsilon; \mathbb{C}) \cong \bigoplus_x S_\kappa(\Gamma_{0,x}(\mathfrak{N}), \varepsilon; \mathbb{C})$ if $Y_0^D(\mathfrak{N}) = \bigsqcup_x \Gamma_{0,x}(\mathfrak{N})\backslash\mathfrak{H}$ follows by the computation similar to the one done just below (4.71) using (SB1–2) in place of (SA1–3). The last isomorphism in (4.95) is a general non-sense (see [LFE] Appendix). The first isomorphism in (4.95) sends $f \in S_\kappa(\Gamma_{0,x}(\mathfrak{N}), \varepsilon; \mathbb{C})$ to a 1-cocycle $c_f(\gamma) = \int_z^{\gamma(z)} f(z)(X_\sigma + zY_\sigma)^{n_\sigma} dz \in L(\kappa^*\varepsilon; \mathbb{C})$ and $f \in \overline{S}_\kappa(\Gamma_{0,x}(\mathfrak{N}), \overline{\varepsilon}; \mathbb{C})$ to a 1-cocycle $c_f(\gamma) = \int_z^{\gamma(z)} f(z)(X_\sigma + \overline{z}Y_\sigma)^{n_\sigma} dz \in L(\kappa^*\varepsilon; \mathbb{C})$, where $I_D = \{\sigma\}$. A detailed proof of this fact is given in [IAT] Chapter 8 (see also [CPS] I, [62a] and [63c] and [LFE] Chapter 6); so, we omit details of the proof of the following theorem.

Theorem 4.36 (Shimura) *Suppose $|I_D| = 1$ and $F \neq \mathbb{Q}$. Then we have a canonical isomorphism of $R(S_0(\mathfrak{N}), \Delta_0(\mathfrak{N}))$-modules,*

$$S_\kappa^D(\mathfrak{N}, \varepsilon; \mathbb{C}) \oplus \overline{S}_\kappa^D(\mathfrak{N}, \overline{\varepsilon}; \mathbb{C}) \cong H^0(Y_0^D(\mathfrak{N}), L(\kappa^*\varepsilon; \mathbb{C})),$$

where $\overline{S}_\kappa^D(\mathfrak{N}, \overline{\varepsilon}; \mathbb{C})$ is made up of complex conjugates of elements in $S_\kappa^D(\mathfrak{N}, \overline{\varepsilon}; \mathbb{C})$ with complex conjugate character $\overline{\varepsilon}$.

4.3.5 Constant Dimensionality

Recall the Neben character $\varepsilon = (\varepsilon_1, \varepsilon_2, \widehat{\varepsilon}_+)$ with values in $W^\times$ as in (ex1). Here $\widehat{\varepsilon}_+(z) = \varepsilon_+(z^{(p)}) z_p^{-\kappa_1-\kappa_2+I}$ for $z \in Z(\mathbb{A}^{(\infty)})$. We define $\widetilde{\varepsilon} = (\widetilde{\varepsilon}_1, \widetilde{\varepsilon}_2, \widetilde{\varepsilon}_+)$ with values in $W^\times$ by the Teichmüller lift of the reduction modulo p of the characters ε. We prove the following theorem.

Theorem 4.37 *If $\kappa_2 - \kappa_1 \geq I$, then $\dim S_\kappa^{D,n.ord}(p\mathfrak{N}, \varepsilon; \mathbb{C})$ depends only on $\widetilde{\varepsilon}$ and $d(D)$ and $\mathfrak{N}$. Moreover if $\kappa_2 - \kappa_1 \geq 2I$, we have*

$$\operatorname{rank} S_\kappa^{M_2(F),n.ord}(p\mathfrak{N}, \varepsilon; W) = \operatorname{rank} S_\kappa^{M_2(F),n.ord}(\mathfrak{N}, \varepsilon; W)$$

for $\mathfrak{N}$ prime to p.

If the assertion of the theorem holds for $\mathfrak{N}$ containing a sufficiently large integer N, the assertion holds for all $\mathfrak{N}$, because by the theory of primitive forms [MFM] 4.6 and [H88] Section 3 (or the strong multiplicity one theorem [AAG] Sections 4 and 10), one can recover the dimension of each space with lower level. Here we give a proof under the assumptions either that $p \geq 5$ and p is unramified in $F/\mathbb{Q}$ or that $Y_0^D(\mathfrak{N})$ is smooth and the sheaf $L(\kappa^*\varepsilon; R)$ is well-defined over $Y_0^D(\mathfrak{N})$. The latter assumptions are valid for $\mathfrak{N}$ having a sufficiently large integer N.

Proof. We choose an open compact subgroup $K \subset S_0(\mathfrak{N})$ so that $Y^D(K)$ is a smooth manifold and that the sheaf $L(\kappa^*\varepsilon; R)$ is well-defined over $Y^D(K)$. We write $\mathcal{Y} = Y^D(K)$ and $\mathcal{Y}' = Y^D(K \cap S_0(p))$. When $Y_0^D(\mathfrak{N})$ is a smooth manifold and $L(\kappa^*\varepsilon; R)$ is well-defined over $Y_0^D(\mathfrak{N})$, we take $K = S_0^D(\mathfrak{N})$. By the same argument as in the proof of Theorem 4.9, the evaluation at $(1, 0)$ of polynomials in $L(\kappa^*\varepsilon; \mathbb{F})$ gives rise to a morphism $L(\kappa^*\varepsilon; \mathbb{F}) \to L(\widetilde{\kappa}\widetilde{\varepsilon}, \mathbb{F})$ of sheaves on $\mathcal{Y}'$, where $\widetilde{\kappa} = (0, I)$ and $\widetilde{\varepsilon}$ is as above. Composing with the restriction map, again by the same argument as in the proof of Theorem 4.9, we get the following isomorphism, if $\kappa_2 - \kappa_1 \geq 2I$,

$$H^q_{n.ord}(\mathcal{Y}, L(\kappa^*\varepsilon; \mathbb{F})) \cong H^q_{n.ord}(\mathcal{Y}', L(\kappa^*\varepsilon; \mathbb{F})) \cong H^q_{n.ord}(\mathcal{Y}', L(\widetilde{\kappa}\widetilde{\varepsilon}; \mathbb{F})).$$

If $\kappa_2 - \kappa_1 > I$, we only have $H^q_{n.ord}(\mathcal{Y}', L(\kappa^*\varepsilon; \mathbb{F})) \cong H^q_{n.ord}(\mathcal{Y}', L(\widetilde{\kappa}\widetilde{\varepsilon}; \mathbb{F}))$. If $\kappa_2 - \kappa_1 = I$, $L(\kappa^*\varepsilon; \mathbb{F}) = L(\widetilde{\kappa}\widetilde{\varepsilon}; \mathbb{F})$ (although $L(\kappa^*\varepsilon; W)$ might be different from $L(\widetilde{\kappa}\widetilde{\varepsilon}; W)$), and we have $H^q_{n.ord}(\mathcal{Y}', L(\kappa^*\varepsilon; \mathbb{F})) \cong H^q_{n.ord}(\mathcal{Y}', L(\widetilde{\kappa}\widetilde{\varepsilon}; \mathbb{F}))$. Thus we need to choose K so that the dimension $\dim S_\kappa^{D,n.ord}(\mathfrak{N}, \varepsilon; \mathbb{C})$ only depends on $\dim_{\mathbb{F}} H^q_{n.ord}(\mathcal{Y}, L(\kappa^*\varepsilon; \mathbb{F}))$. We specify K later.

By Theorem 4.34, $\dim_{\mathbb{C}} S_\kappa^D(\mathfrak{N},\varepsilon;\mathbb{C})$ is independent of D as long as the quadruple $(d(D),\kappa,\varepsilon,\mathfrak{N})$ is fixed; so, we may assume that $q = |I_D| \le 1$, because, for a given finite set of primes S, we can always find a quaternion algebra D with $d(D) = \prod_{\mathfrak{q}\in S}\mathfrak{q}$ and $g - |I_D| = |I^D| \equiv |S| \mod 2$ by a Hasse principle (see [IAT] (9.2.4–6) and [BNT]). Since the projectors e and e_0 are defined taking the p-adic limit, they give idempotents of $h_\kappa(\mathfrak{N},\varepsilon;W)$. However, every idempotent of $h_\kappa(\mathfrak{N},\varepsilon;W)$ comes from an idempotent of $h_\kappa(\mathfrak{N},\varepsilon;\mathcal{W})$ if $\mathcal{W}$ is sufficiently large, since $h_\kappa(\mathfrak{N},\varepsilon;W) = h_\kappa(\mathfrak{N},\varepsilon;\mathcal{W})\otimes_{\mathcal{W}} W$. Thus these idempotents are well-defined over $\mathcal{W}$ and over $\mathbb{C}$; so, $H^q_{n.ord} = e\cdot H^q_{n.ord}$ makes sense for any coefficient ring.

By definition, $S_\kappa^D(\mathfrak{N},\varepsilon;R)$ is a quotient of $H^0(Y_0^D(\mathfrak{N}),L(\kappa^*\varepsilon;R))$ when $q=0$, and as we have seen in (4.92),

$$S_\kappa^{D,n.ord}(\mathfrak{N},\varepsilon;R) \cong H^0_{n.ord}(Y_0^D(\mathfrak{N}),L(\kappa^*\varepsilon;R)) \text{ for all } \kappa \text{ and } \varepsilon \text{ if } q=0.$$

Similarly, by Theorem 4.36, we have

$$S_\kappa^D(\mathfrak{N},\varepsilon;\mathbb{C})^2 \cong H^0(Y_0^D(\mathfrak{N}),L(\kappa^*\varepsilon;\mathbb{C})) \text{ for all } \kappa \text{ and } \varepsilon \text{ if } q=1.$$

Since by the universal coefficient theorem (e.g., [SFT] Theorem II.18.3), we have from the flatness of $\mathbb{C}$ and W over $\mathcal{W}$ that

$$H^q_{n.ord}(Y_0^D(\mathfrak{N}),L(\kappa^*\varepsilon;\mathbb{C})) = H^q_{n.ord}(Y_0^D(\mathfrak{N}),L(\kappa^*\varepsilon;\mathcal{W}))\otimes_{\mathcal{W}}\mathbb{C},$$
$$H^q_{n.ord}(Y_0^D(\mathfrak{N}),L(\kappa^*\varepsilon;W)) = H^q_{n.ord}(Y_0^D(\mathfrak{N}),L(\kappa^*\varepsilon;\mathcal{W}))\otimes_{\mathcal{W}} W.$$

Thus we need to prove the constancy of $\operatorname{rank}_W H^q_{n.ord}(Y_0^D(\mathfrak{N}),L(\kappa^*\varepsilon;W))$ and that $H^q_{n.ord}(Y_0^D(\mathfrak{N}),L(\kappa^*\varepsilon;W))$ is W-free. Take a prime $\mathfrak{l}\nmid p\mathfrak{N}$ such that $p\nmid (N(\mathfrak{l})^2-1)$. As we have already remarked in the proof of Theorem 4.10, there exist infinitely many such primes $\mathfrak{l}$ if $p\ge 5$ and p is unramified in $F/\mathbb{Q}$. We take $K = S_0(\mathfrak{N})\cap S_1^1(\mathfrak{l})$. Then $H^q(Y_0^D(\mathfrak{N}),L(\kappa^*\varepsilon;R))$ is canonically a direct summand (as Hecke modules) of $H^q(\mathcal{Y},L(\kappa^*\varepsilon;R))$ by the restriction map for all $\mathcal{W}$-algebras R, because the restriction followed by the transfer is a multiplication by an integer factor of $[S_0(\mathfrak{N}),K]\big|(N(\mathfrak{l})^2-1)$ (cf. [MFG] Proposition 4.21), which is invertible in R. Thus we only need to prove these facts for $H^q_{n.ord}(\mathcal{Y},L(\kappa^*\varepsilon;W))$.

We may further choose $\mathfrak{l}$ so that the image $\overline{\Gamma}_x$ of $\Gamma_x = x\cdot Kx^{-1}\cap G_D(\mathbb{Q})$ in $G_D^{ad}(\mathbb{Q}) = G_D(\mathbb{Q})/Z(\mathbb{Q})$ is torsion-free for all $x\in G_D(\mathbb{A}^{(\infty)})$. Indeed, if $\gamma\in\Gamma_x$ is torsion in $\overline{\Gamma}_x$, $F[\gamma]$ is a quadratic extension of F with $\gamma^N=\epsilon\in O^\times$ for the order N of the image of γ in $\overline{\Gamma}_x$. Write M for the extension of F obtained by adding $\sqrt{u}$ for all $u\in O^\times$. Since $[O^\times:(O^\times)^2]=2^g$, $[M:F]\le 2^g$. In M, $\zeta=\gamma\sqrt{\epsilon}^{-1}$ has finite order N; so, ζ is a root of unity in $M[\gamma]$. Since $\epsilon\equiv 1 \mod \mathfrak{l}$, we may assume that $\zeta\equiv 1 \mod \mathfrak{L}$ for a prime factor $\mathfrak{L}$ of $\mathfrak{l}$ in $M[\gamma]$. This shows either $\zeta=1$ or $\ell|N$ for the residual characteristic ℓ of $\mathfrak{l}$. This never happens if ℓ is sufficiently large, because N is bounded. If $\zeta=1$, then γ has two eigenvalues $\pm\sqrt{\epsilon}$; so, $\sqrt{\epsilon}\equiv -\sqrt{\epsilon}\equiv 1 \mod \mathfrak{l}$, which happens only if $\ell=2$. Thus we have plenty of choice of $\mathfrak{l}$ as above. Under the above

choice, $\overline{\Gamma}_x \subset Z(\mathbb{Z})$ if $q = 0$ and $\mathcal{Y}$ is smooth if $q = 1$. Since $Z(\mathbb{Z}) = O^\times$ acts on $L(\kappa^*\varepsilon; R)$ trivially, $L(\kappa^*\varepsilon; R)$ is a well-defined sheaf over $\mathcal{Y}$.

For simplicity, we write $H^q_{n.ord}(\mathcal{Y}, L(\kappa^*\varepsilon; R))$ as $H^q_{n.ord}(L(\kappa^*\varepsilon; R))$. By the long exact sequence attached to $L(\kappa^*\varepsilon; W) \overset{p^m}{\hookrightarrow} L(\kappa^*\varepsilon; W) \twoheadrightarrow L(\kappa^*\varepsilon; \mathbb{F})$ for the residue field $\mathbb{F}$ of W, we have another exact sequence:

$$H^q_{n.ord}(L(\kappa^*\varepsilon; W)) \otimes_W W_m \hookrightarrow H^q_{n.ord}(L(\kappa^*\varepsilon; W_m)) \twoheadrightarrow H^{q+1}_{n.ord}(L(\kappa^*\varepsilon; W))[p^m].$$

When $q = 0$, $H^1_{n.ord}(L(\kappa^*\varepsilon; W)) = 0$ because $\dim \mathcal{Y} = 0$. Thus we get

$$H^0_{n.ord}(L(\kappa^*\varepsilon; W)) \otimes_W W_m \cong H^0_{n.ord}(L(\kappa^*\varepsilon; W_m)),$$

and $H^0_{n.ord}(L(\kappa^*\varepsilon; W))$ is W-free.

Suppose $q = 1$. Then $H^2(L(\kappa^*\varepsilon; W))$ is the dual of $H^0(L(\widehat{\kappa}\varepsilon^{-1}; W^*))$ by Poincaré duality, where W^* is the Pontryagin dual $\bigcup_n p^{-n}W/W$ of W. Here the adjoint operator of $[KgK]$ acting on $H^0(L(\widehat{\kappa}\varepsilon^{-1}; W^*))$ is given by $[Kg^{-\iota}K] = [K \cdot N(g)^{-1}K][KgK]$ for the reduced norm map $N : G_D \to Z$. Since the action of the center $[K \cdot N(g)K]$ is the multiplication by a p-adic unit, we have the duality between the nearly ordinary part of the two cohomology groups $H^2_{n.ord}(L(\kappa^*\varepsilon; W))$ and $H^0_{n.ord}(L(\widehat{\kappa}\varepsilon^{-1}; W^*))$.

We now show $H^0_{n.ord}(L(\widehat{\kappa}\varepsilon^{-1}; W^*)) = 0$. Since we have

$$H^0_{n.ord}(L(\widehat{\kappa}\varepsilon^{-1}; W^*)) = \bigcup_m H^0_{n.ord}(L(\widehat{\kappa}\varepsilon^{-1}; p^{-m}W/W)),$$

we only need to show that $T(p)$ or $U(p)$ is nilpotent on $H^0(L(\widehat{\kappa}\varepsilon^{-1}; W_1))$. We regard $H^0(L(\widehat{\kappa}\varepsilon^{-1}; W_1))$ as a space of locally constant functions on $G_D(\mathbb{A})$ with values in $L(\widehat{\kappa}; W_1)$ satisfying the automorphy property as in (SB). Choose an integer $n > 0$ so that $p^n \equiv 1 \mod \mathfrak{l}$. Decomposing $t = \mathcal{T}(p^n) = \bigsqcup_h hK$, we have from (4.90) that $f|t(x) = p^{-n\widehat{\kappa}_1} \sum_h h \cdot f(xh^{-\iota})$. We may assume $\det(h) = \varpi_{\mathfrak{p}}$ and that h is of the form $\left(\begin{smallmatrix} \alpha & u \\ 0 & \beta \end{smallmatrix}\right)$ with $\alpha\beta = p^n$ for u running over representatives modulo αO_p. By the strong approximation theorem, connected components of $\mathcal{Y}$ are indexed by $T(\mathbb{A})/N(K)T(F)T(\mathbb{R})^+ = Cl_F(\mathfrak{l})$ for the reduced norm map $N : G_D \to T$. Since f is constant on each connected component of $\mathcal{Y}$ and $T(p^n)$ or $U(p^n)$ does not move the connected component (i.e., because the ideal (p^n) is generated by a totally positive element $p^n \equiv 1 \mod \mathfrak{l}$ in F), $f(xh^{-\iota})$ is a polynomial $P(X, Y)$ independent of h.

By the definition of the action, for $P_h(X, Y) = f(xh^{-\iota})$, we have $p^{-n\widehat{\kappa}_1} h \cdot P_h(X, Y) = P(\beta X - uY, \alpha Y)$ for $r + s = n$. Therefore writing P as a linear combination of $\prod_\sigma P_\sigma(X_\sigma, Y_\sigma)$, we compute $h \cdot P$ at each piece $P_\sigma(X_\sigma, Y_\sigma)$. Writing $P_u(X_\sigma, Y_\sigma) = P_\sigma(p^r X_s - uY_\sigma, p^s Y_s) = P_\sigma(p^r X_s - uY_\sigma, 0)$ if $s > 0$ and $P'(X_\sigma, Y_\sigma) = P_\sigma(0, Y_s)$, we see $\sum_u P_\sigma(p^r X_\sigma - uY_\sigma, 0) = 0$ in $L(\widehat{\kappa}\varepsilon^{-1}; W_1)$. For $U(\mathfrak{p})$, only those $P_\sigma(p^r X_\sigma - uY_\sigma, 0)$ appear; so, $U(\mathfrak{p}) = 0$ on the cohomology. Since $U(p) = 0$, after applying $T(p)$, we only get a monomial $P_\sigma(0, Y_s)$ of Y_σ. When we consider $T(p)$, p is prime to $\mathfrak{N}$. So $P(X, Y)$ is invariant under $U_0(\mathfrak{N})_p = GL_2(O_p)$, but a nonzero monomial of Y_σ is never

invariant under $GL_2(O_p)$; so, $P_\sigma(X_\sigma, Y_\sigma)$ does not involve a monomial of Y_σ, that is, $P_\sigma(0, Y_\sigma) = 0$. This shows the desired vanishing. In this proof of $H^2_{n.ord}(\mathcal{Y}, L(\kappa^*\varepsilon; R)) = 0$, we have used the strong approximation theorem to show that $T(p^n)$ and $U(p^n)$ for a suitable $n > 0$ preserve each connected component of $\mathcal{Y}$; so, obviously the argument does not work if $I_D = \emptyset$ where the strong approximation theorem for $G_{D,1}(\mathbb{A}^{(\infty)})$ (for the derived subgroup $G_{D,1} = \mathrm{Ker}(N : G_D \to T) \subset G_D$) does not hold.

We find $H^1_{n.ord}(L(\kappa^*\varepsilon; W)) \otimes_W W_m = H^1_{n.ord}(L(\kappa^*\varepsilon; W_m))$ by the vanishing of $H^2_{n.ord}(L(\kappa^*\varepsilon; W))$. Again by the exact sequence

$$H^0_{n.ord}(L(\kappa^*\varepsilon; W)) \otimes_W W_m \hookrightarrow H^0_{n.ord}(L(\kappa^*\varepsilon; W_m)) \twoheadrightarrow H^1_{n.ord}(L(\kappa^*\varepsilon; W))[p^m]$$

and the vanishing of $H^0_{n.ord}$, we find that $H^1_{n.ord}(L(\kappa^*\varepsilon; W))$ is torsion-free if $q = 1$. This finishes the proof.

4.4 Appendix: Fundamental Groups

This section gives a sketch of the theory of algebraic fundamental groups which has been and will be used in the proof of the global reciprocity laws (for example, Theorem 4.14 and Theorem 6.26). At the end, we summarize group-theoretic tools also used in the proof of the reciprocity laws.

4.4.1 Categorical Galois Theory

We give a brief exposition of how one can axiomatize Galois theory (the idea is due to Grothendieck, and the main reference is [SGA] V.4). We assume a minimal knowledge of category theory and terminology which one can find in [HAL] Chapter II, [GME] Sections 4–5, and [MFG] Section 4. This subsection interprets the result in Section 2.3.1 in terms of category theory.

We recall some terminology in category theory. Let $\mathcal{C}$ be a category. A morphism $X \xrightarrow{f} Y$ in this category is called an *epimorphism* if for any third object Z and morphisms $Y \underset{h}{\overset{g}{\rightrightarrows}} Z$, $g \circ f = h \circ f \Leftrightarrow g = h$ (so f is a right unit). In other words, for any other object Z in $\mathcal{C}$, the pullback map $\mathrm{Hom}_{\mathcal{C}}(Y, Z) \xrightarrow{f^*} \mathrm{Hom}_{\mathcal{C}}(X, Z)$ is injective if f is an epimorphism. Similarly a morphism $X \xrightarrow{f} Y$ in this category is called a *monomorphism* if for any third object Z and morphisms $Z \underset{h}{\overset{g}{\rightrightarrows}} X$, $f \circ g = f \circ h \Leftrightarrow g = h$ (so f is a left unit).

Let $SETS$ be the category of sets. Recall that a covariant (resp. contravariant) functor $F : \mathcal{C} \to SETS$ is called *representable* if there is an object $\mathcal{F}$ in $\mathcal{C}$ such that we have a system of functorial bijections $\phi_X : F(X) \to \mathrm{Hom}_{\mathcal{C}}(\mathcal{F}, X)$ (resp. $\phi_X : F(X) \to \mathrm{Hom}_{\mathcal{C}}(X, \mathcal{F})$). Here the word "functorial" means that $f_* \circ \phi_X = \phi_Y \circ F(f)$ in the covariant case and $f^* \circ \phi_Y = \phi_X \circ F(f)$ in the contravariant case for any morphism $X \xrightarrow{f} Y$ in $\mathcal{C}$, where $f_*(g) = f \circ g$ for

$g \in \mathrm{Hom}_{\mathcal{C}}(\mathcal{F}, X)$ and $f^*(h) = h \circ f$ for $h \in \mathrm{Hom}_{\mathcal{C}}(Y, \mathcal{F})$. The object in $\mathcal{C}$ representing a functor is unique up to isomorphisms (a characterization by a universal property; see Section 3.1.5 in the text and [GME] Lemma 4.1).

If a covariant functor $F_{\mathbf{0}}$ sending every object in $\mathcal{C}$ to a singleton $\{\mathbf{0}\}$ is representable in $\mathcal{C}$, the object representing the functor is called an *initial* object in $\mathcal{C}$. Similarly if a contravariant functor $F^{\mathbf{0}}$ sending every object in $\mathcal{C}$ to a singleton $\{\mathbf{0}\}$ is representable in $\mathcal{C}$, the object representing the functor is called a *final* object in $\mathcal{C}$.

If the covariant functor $T \mapsto \mathrm{Hom}_{\mathcal{C}}(X, T) \times \mathrm{Hom}_{\mathcal{C}}(Y, T)$ for two given objects X, Y of $\mathcal{C}$ is representable by an object of $\mathcal{C}$, the object is written as $X \oplus Y$ and is called the *direct sum*, and we have inclusions $i_X : X \to X \oplus Y$ and $i_Y : Y \to X \oplus Y$ such that $\mathrm{Hom}_{\mathcal{C}}(X \oplus Y, T) \cong \mathrm{Hom}_{\mathcal{C}}(X, T) \times \mathrm{Hom}_{\mathcal{C}}(Y, T)$ is given by $\phi \mapsto (\phi \circ i_X, \phi \circ i_Y)$ (see [MFG] 4.4). If the contravariant functor: $T \mapsto \mathrm{Hom}_{\mathcal{C}}(T, X) \times \mathrm{Hom}_{\mathcal{C}}(T, Y)$ for two given objects X, Y of $\mathcal{C}$ is representable by an object of $\mathcal{C}$, the object is written as $X \times Y$ and is called a *direct product*, and we have projections $p_X : X \times Y \to X$ and $p_Y : X \times Y \to Y$ such that the isomorphism $\mathrm{Hom}_{\mathcal{C}}(T, X \times Y) \cong \mathrm{Hom}_{\mathcal{C}}(T, X) \times \mathrm{Hom}_{\mathcal{C}}(T, Y)$ is given by $\phi \mapsto (p_X \circ \phi, p_Y \circ \phi)$ (see [MFG] 4.4).

For a given object S in $\mathcal{C}$, we write $\mathcal{C}_{/S}$ for the category of arrows $X \xrightarrow{f} S$. Morphisms from $X \xrightarrow{f} S$ into $Y \xrightarrow{g} S$ are given by morphisms $h : X \to Y$ with $g \circ h = f$. If we apply the above construction to $\mathcal{C}_{/S}$, the direct product in $\mathcal{C}_{/S}$ if it exists is written as $X \times_S Y$ and is called the *fibered product* of X and Y over S (see [MFG] 4.2.1). So in this case, the contravariant functor $T \mapsto \mathrm{Hom}_{\mathcal{C}_{/S}}(T, X) \times \mathrm{Hom}_{\mathcal{C}_{/S}}(T, Y)$ is representable by $X \times_S Y$ in $\mathcal{C}_{/S}$ (see [GME] 5.2). If $\mathcal{C} = SETS$, we see easily that

$$X \times_S Y = \left\{(x, y) \in X \times Y \middle| f(x) = g(y)\right\}.$$

Dualizing the above construction, for an object S of $\mathcal{C}$, we consider a category $\mathcal{C}^S$ whose objects are arrows $S \xrightarrow{f} X$ and $\phi \in \mathrm{Hom}(S \xrightarrow{f} X, S \xrightarrow{g} Y)$ is a morphism $\phi : X \to Y$ with $\phi \circ f = g$. Then $X \oplus^S Y$ in $\mathcal{C}^S$ if it exists represents the covariant functor $T \mapsto \mathrm{Hom}_{\mathcal{C}^S}(X, T) \times \mathrm{Hom}_{\mathcal{C}^S}(Y, T)$. The object $X \oplus^S Y$ is called the *fibered sum* of X and Y under S. If $\mathcal{C} = SETS$, we see easily that $X \oplus^S Y$ is the quotient of the disjoint union $X \sqcup Y$ under the relation that $x \sim y \Leftrightarrow x = f(s)$ and $y = g(s)$ (see [MFG] 4.2.1).

An epimorphism $X \xrightarrow{f} Y$ in $\mathcal{C}$ is called a *strict* epimorphism if for any other object Z of $\mathcal{C}$

$$\mathrm{Im}(f^* : \mathrm{Hom}_{\mathcal{C}}(Y, Z) \to \mathrm{Hom}_{\mathcal{C}}(X, Z)) = \left\{g \in \mathrm{Hom}_{\mathcal{C}}(X, Z) \middle| p_1^* g = p_2^* g\right\},$$

where $p_j : X \times_Y X \to X$ are the two projections. If this is the case, we call the following sequence *exact*: $X \times_Y X \overset{p_1, p_2}{\rightrightarrows} X \xrightarrow{f} Y$. The notion of strict epimorphisms is stronger than that of epimorphisms. For example, $i : Spec(\mathbb{Q}) \to Spec(\mathbb{Z})$ is an epimorphism in SCH (because it is dominant).

Since $\mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Q} = \mathbb{Q}$, if i were a strict epimorphism, the empty set $\mathrm{Hom}_{\mathrm{alg}}(\mathbb{Q}, \mathbb{Z})$ has to be identical to the singleton $\mathrm{Hom}_{\mathrm{alg}}(\mathbb{Q}, \mathbb{Q})$. This is a contradiction, and hence i is not a strict epimorphism. A faithfully flat morphism $X \to Y$ of finite type is a strict epimorphism (see [ECH] Theorem I.2.17) in the category of schemes. In the category $SETS$, the notion of strict epimorphisms and that of epimorphisms coincide and are equivalent to surjectivity.

Let $G \subset \mathrm{Aut}_{\mathcal{C}}(X)$ be a finite subgroup for an object X of $\mathcal{C}$. Write $\mathrm{Hom}_{\mathcal{C}}(X,T)^G$ for the set of morphisms $X \xrightarrow{f} T$ such that $f \circ g = f$ for all $g \in G$. If the covariant functor $T \mapsto \mathrm{Hom}_{\mathcal{C}}(X,T)^G$ is representable by an object Y of $\mathcal{C}$ with a unique morphism $\pi : X \to Y$ inducing $\pi^* : \mathrm{Hom}_{\mathcal{C}}(Y,T) \cong \mathrm{Hom}_{\mathcal{C}}(X,T)^G$, we call Y the *categorical* quotient of X by G and the morphism $\pi : X \to Y$ the quotient morphism (see [GME] 8.2).

Here is an example of $\mathcal{C}$ with fibered products and fibered sums. We start with a profinite compact group Π. A typical example for Π is the Galois group $\mathrm{Gal}(K/k)$ of a (possibly infinite) Galois extension K/k. We denote $\mathcal{C}(\Pi)$ for the category of finite sets on which Π acts continuously (from the left). Thus for each object X in $\mathcal{C}(\Pi)$, the stabilizer in Π of each point $x \in X$ is an open subgroup of Π. For X, Y in $\mathcal{C}(\Pi)$, a map $f : X \to Y$ is a morphism of $\mathcal{C}(\Pi)$ if $f(\pi(x)) = \pi(f(x))$ for all $x \in X$ and $\pi \in \Pi$. The collection of all such maps is written as $\mathrm{Hom}_{\Pi}(X,Y)$. Let $\mathbf{0}$ be the set of a single point on which Π acts trivially. Then $\mathrm{Hom}_{\Pi}(X, \mathbf{0})$ is made of one map taking everything in X to $\mathbf{0}$. In other words, $\mathbf{0}$ is the final object. The empty set $\emptyset$ is considered to be an object in $\mathcal{C}(\Pi)$, which gives an initial object in $\mathcal{C}(\Pi)$. Since $\mathcal{C}(\Pi)$ is a subcategory of $SETS$, the fibered product $X \times_S Y$ for $X \xrightarrow{f} S$ and $Y \xrightarrow{g} S$ is given by $X \times_S Y = \{(x,y) \in X \times Y \,|\, f(x) = g(y)\}$ with two projections p_X and p_Y induced by two projections $X \times Y$ onto X and Y. The fibered sum of $S \xrightarrow{f} X$ and $S \xrightarrow{g} Y$ is given by the set-theoretic fibered sum $X \oplus^S Y$ already described, and in particular, the direct sum is the fibered sum for an initial object S and is $X \sqcup Y$ with two obvious inclusions i_X and i_Y. If $G \subset \mathrm{Aut}_{\mathcal{C}(\Pi)}(X)$ is a finite group, the set-theoretic quotient X/G gives the categorical quotient in $\mathcal{C}(\Pi)$. If $X \xrightarrow{f} Y$ is a morphism in $\mathcal{C}(\Pi)$, it factors as $X \twoheadrightarrow \mathrm{Im}(f) \hookrightarrow Y$. Here $p : X \twoheadrightarrow \mathrm{Im}(f)$ is a surjection by definition and hence is a strict epimorphism, and $i : \mathrm{Im}(f) \hookrightarrow Y$ is a monomorphism inducing a direct sum decomposition $Y = (Y - \mathrm{Im}(f)) \sqcup \mathrm{Im}(f)$. Thus we have verified the following three properties for $\mathcal{C} = \mathcal{C}(\Pi)$.

(C1) $\mathcal{C}$ has a final object and a fibered product of any two morphisms $X \to S$ and $Y \to S$ in $\mathcal{C}$ over S.

(C2) $\mathcal{C}$ has an initial object and a direct sum of any two objects in $\mathcal{C}$. For each object X of $\mathcal{C}$ and each finite subgroup G of $\mathrm{Aut}_{\mathcal{C}}(X)$, the categorical quotient X/G exists.

(C3) Each morphism $X \xrightarrow{f} Y$ in $\mathcal{C}$ factors as $X \xrightarrow{f'} X' \xrightarrow{f''} Y$ so that f' is a strict epimorphism and f'' is a monomorphism.

A category $\mathcal{C}$ satisfying the three conditions (C1–3) is called a *Galois* category. A covariant functor $F : \mathcal{C} \to \mathcal{C}'$ is called *left exact* if it takes monomorphisms of $\mathcal{C}$ to monomorphisms of $\mathcal{C}'$ and $F(X \times_S Y) = F(X) \times_{F(S)} F(Y)$. We note that $F(X) \times_{F(S)} F(Y) = \{(x, y) \in F(X) \times F(Y) | F(f)(x) = F(g)(y)\}$ for $X \xrightarrow{f} S$ and $Y \xrightarrow{g} S$ if $\mathcal{C}' = SETS$. Similarly a covariant functor $F : \mathcal{C} \to \mathcal{C}'$ is called *right exact* if it takes epimorphisms of $\mathcal{C}$ to epimorphisms of $\mathcal{C}'$ and commutes with fibered sums. A left exact covariant functor $F : \mathcal{C} \to \mathcal{C}'$ is called *exact* if it preserves monomorphisms, epimorphisms, and exact sequences. The inclusion functor $F : \mathcal{C}(\Pi) \hookrightarrow SETS$ obviously satisfies the following three conditions:

(F1) F is left exact and has values in the full subcategory of finite sets in $SETS$;

(F2) $F(X \oplus Y) = F(X) \sqcup F(Y)$ for objects X, Y in $\mathcal{C}$; $F(f)$ is a surjection if $F : X \to Y$ is a strict epimorphism; and if $X \xrightarrow{\pi} Y$ is the categorical quotient in $\mathcal{C}$ by a finite group $G \subset \mathrm{Aut}_{\mathcal{C}}(X)$, $F(X) \xrightarrow{F(\pi)} F(Y)$ is a categorical quotient in $SETS$ by the same group G sitting inside $\mathrm{Aut}(F(X))$, which is just a set-theoretic quotient;

(F3) a morphism $X \xrightarrow{f} Y$ in $\mathcal{C}$ is an isomorphism if $F(f) : F(X) \to F(Y)$ is an isomorphism.

For a Galois category $\mathcal{C}$, a covariant functor $F : \mathcal{C} \to SETS$ satisfying the three conditions (F1-3) is called a *fundamental* functor.

We quote the following theorem of Grothendieck from [SGA] V.4.

Theorem 4.38 *If a pair $(\mathcal{C}, F)$ of a category $\mathcal{C}$ and a covariant functor $F : \mathcal{C} \to SETS$ satisfies the six conditions* (C1–3) *and* (F1–3)*, there exists a unique profinite group Π (up to isomorphisms) such that $\mathcal{C}$ is equivalent to the category $\mathcal{C}(\Pi)$.*

The group Π is called the fundamental group of $(\mathcal{C}, F)$ and sometimes written as $\pi_1(\mathcal{C}, F)$. We refer to [SGA] V.4 for the proof. We call an object X in $\mathcal{C}$ *connected* if it is not a direct sum of two non initial objects. Here we list steps of the proof and give a brief description of each step.

1. A morphism $f \in \mathrm{Hom}_{\mathcal{C}}(X, Y)$ is a monomorphism if and only if $F(f)$ is an injection.

   In fact $p_X : X \times_Y X \to X$ is an isomorphism if and only if $f : X \to Y$ is a monomorphism. Indeed, if $h, g \in \mathrm{Hom}_{\mathcal{C}}(Z, X)$ satisfy $f \circ g = f \circ h = f'$, regarding Z as an object in $\mathcal{C}_{/Y}$ by f', we have $g \times_Y h \in \mathrm{Hom}(Z, X \times_Y X)$. Thus f being a monomorphism is equivalent to the fact that the two functors $Z \mapsto \mathrm{Hom}_{\mathcal{C}}(Z, X)$ and $Z \mapsto \mathrm{Hom}_{\mathcal{C}}(Z, X \times_Y X)$ are isomorphic. This implies by the uniqueness of the object representing a functor (see [GME] Lemma 4.1) the above statement that $X \times_Y X \cong X$. Then by (C1), (F1), and (F3) combined, we conclude the assertion. □

2. A sequence $X_1 \xrightarrow{f_1} X_2 \xrightarrow{f_2} X_3 \to \cdots \to X_j \to X$ of monomorphisms f_j for an object X in $\mathcal{C}$ stabilizes after the finite steps; that is, for sufficiently large j, f_j are all isomorphisms.

   This assertion is obvious in the category of finite sets. Since F preserves monomorphisms and brings $\mathcal{C}$ into the category of finite sets, it follows by (F3) from the result for the category of finite sets. □

3. There is a connected partially ordered set I and a projective system $\{P_i, \pi_{ij} : P_i \to P_j \ (i > j)\}_{i \in I}$ in $\mathcal{C}$ such that
   a) for each object X in $\mathcal{C}$ we have a unique morphism $\phi_i : P_i \to X$ for each sufficiently large $i \in I$ compatible with the projective system;
   b) there exist a universal element $\varphi = \varprojlim_i \varphi_i \in \varprojlim_{i \in I} F(P_i)$ and an isomorphism $\iota_X : F(X) \cong \varinjlim_{i \in I} \mathrm{Hom}_{\mathcal{C}}(P_i, X)$ induced by $\iota_X(f) = \varprojlim_i F(\phi_i)(\varphi_i)$.

   We hereafter write $\mathrm{Hom}(P, X)$ for $\varinjlim_{i \in I} \mathrm{Hom}_{\mathcal{C}}(P_i, X)$, and consider $P = \varprojlim_{i \in I} P_i$ in the procategory of $\mathcal{C}$. The object P is called the *fundamental pro-object* of $\mathcal{C}$ with respect to the fundamental functor F.

4. P_i can be chosen so that
   a) P_i for all i is connected;
   b) P_i is not an initial object for all i;
   c) π_{ij} is an epimorphism for all $i > j$.

   For an axiomatic proof of (3) and (4) which is based on the assertions (1) and (2) besides the axioms (C1–3) and (F1–3), see [FGA] Exposé 195.

   We describe what is going on by examples. Suppose that $\mathcal{C}$ is the opposite category of the category of separable finite-dimensional semi-simple commutative algebras over a given field k. In other words, $\mathcal{C}$ is made of $Spec(K)$ for separable semi-simple extensions K/k with $\dim_k K < \infty$. Let Ω (resp. Ω^{sep}) be an algebraic (resp. separable) closure of k ($Spec(\Omega^{sep})$ is a fundamental pro-object). Then $F : Spec(K) \mapsto \mathrm{Hom}_k(K, \Omega^{sep}) = \mathrm{Hom}_{\mathcal{C}}(Spec(\Omega^{sep}), Spec(K))$ gives the covariant functor. One can easily check that $(\mathcal{C}, F)$ is a Galois category. In this case, we may take I to be the set of all finite Galois extensions in the fixed algebraically closed field Ω containing k. The order of I is induced by the inclusion of two fields.

   If $\mathcal{C}$ is the category of finite étale coverings of an irreducible variety S of finite type over a field k, we fix a geometric point $\overline{s} \in S(\Omega)$ and take $F(X) = \mathrm{Hom}_{SCH/S}(\overline{s}, X)$. We often write this category as ET/S. In this case, I is the category of all étale irreducible Galois coverings with ordering $X \to Y$ on I given by S-morphisms. For each $X \in I$, its function field is a finite separable extension of the function field of S. Since the function field determines X over S (because epimorphisms in the category of schemes are dominant morphisms), the category I is actually a set.

5. If X is not an initial object and Y is connected, any morphism $f : X \to Y$ in $\mathcal{C}$ is a strict epimorphism.

   In the above example of the two categories (of semi-simple separable finite-dimensional commutative algebras over k and of étale coverings over S), this assertion is obvious. However this is just a consequence of the axioms (C3) and (F2–3). An axiomatic proof is in [SGA] V.4.e.

6. The following conditions are equivalent.
   a) The natural injection: $\mathrm{Hom}_{\mathcal{C}}(P_i, P_i) \to \mathrm{Hom}(P, P_i) \xrightarrow{\iota_{P_i}} F(P_i)$ is surjective;
   b) $\mathrm{Aut}(P_i)$ acts on $F(P_i)$ transitively;
   c) $\mathrm{Aut}(P_i)$ acts on $F(P_i)$ simply transitively.

   We call P_i *of Galois type* if one of the above equivalent properties is satisfied by P_i.

   This is again obvious for our two examples of $\mathcal{C}$, because $\mathrm{End}(P_i)$ is the monoid of field endomorphisms of a finite extension over a given base field k. To prove this axiomatically, we apply the functor F and prove the fact in the image of F by using the previous assertions (1), (4) and (5).

7. Every object X in $\mathcal{C}$ is covered by an object P_i of Galois type so that the unique map $P \to X$ in 3.a) factors through P_i.

8. Let I^{gal} be the subset of I made of all indices i with P_i of Galois type. Then we have $\mathrm{Hom}(P, P) = \mathrm{Aut}(P) = \varprojlim_{i \in I^{gal}} \mathrm{Aut}_{\mathcal{C}}(P_i)$, which is set-theoretically in bijection with $F(P) = \varprojlim_i F(P_i)$. We put $\Pi = \mathrm{Aut}(P)$.

9. Define a functor $G : \mathcal{C}(\Pi) \to \mathcal{C}$ so that

$$\mathrm{Hom}_{\mathcal{C}}(G(Z), ?) \cong \mathrm{Hom}_{\Pi}(Z, \mathrm{Hom}(P, ?))$$

   as functors with variable X in $\mathcal{C}$.

   For each connected X in $\mathcal{C}(\Pi)$, the group Π acts transitively on X by its connectedness; so, $X \cong \Pi/H$ for a stabilizer of a point $x \in X$. Since H acts faithfully on P, for sufficiently large j in I^{gal}, if $i > j$, π_{ij} induces an isomorphism $P_i/H \cong P_j/H$ of the categorical quotients, which gives P/H (independently of $i > j$). We put $G(X) = P/H$. Here the existence of the categorical quotients P_i/H in $\mathcal{C}$ is guaranteed by (C2). If X is not connected, we can write X uniquely as $X = \bigsqcup_Z Z$ for connected components Z. Then we define $G(X) = \bigoplus_Z G(Z)$ which gives the inverse G of F. See [SGA] V.4.h for more details of the argument.

   In our example of Galois extensions, one can check that $Spec(K)/H = Spec(K^H)$ gives the categorical quotient of $Spec(K)$ by a finite group $H \subset \mathrm{Gal}(K/k)$. For the category of étale coverings of a base scheme S, the existence of categorical quotients is explained in [GME] Proposition 8.4.

10. The two functors F and G give the equivalence of categories.

    In our two examples, this point is plain. □

We now quote two more propositions from [SGA] V Propositions 5.5–6.

Proposition 4.39 *Let $\{Q_i\}_{i\in I}$ be a pro-object in $\mathcal{C} = \mathcal{C}(\Pi)$ (i.e., a projective system of objects in $\mathcal{C}$). Suppose that Q_i is not an initial object for any i and that Q_i is connected for all i. Define a covariant functor $G : \mathcal{C} \to SETS$ by $G(X) = \mathrm{Hom}_{\mathcal{C}}(Q, X)$. Then Q is isomorphic to Π/H for a closed subgroup H of Π, and the functor G is isomorphic to the functor $E \mapsto E^H$ for the fixed point subset E^H of E under H.*

The subgroup H is given by the stabilizer of one point $q \in Q$. Since Q_i is connected, the action of Π is transitive, and $Q \cong \Pi/H$ via $\pi q \leftrightarrow \pi H$. □

Proposition 4.40 *Let $(\mathcal{C}, F, P)$ be as above (a Galois category, a fundamental functor, and the fundamental pro-object with respect to F). Let $P' = \varprojlim_{j\in J} P'_j$ be a pro-object in the procategory of $\mathcal{C}$, and define $F' : \mathcal{C} \to SETS$ by $F'(X) = \mathrm{Hom}(P', X) = \varinjlim_{j\in J} \mathrm{Hom}_{\mathcal{C}}(P'_j, X)$. Then the following three conditions are equivalent.*

1. *$P' \cong P$ and also $F' \cong F$;*
2. *P' is fundamental and also F' is fundamental;*
3. *F' brings direct sums to direct sums, and $F(X) \neq \emptyset$ if X is not an initial object.*

We may identify $\mathcal{C}$ with $\mathcal{C}(\Pi)$. For $\mathcal{C}(\Pi)$, the fundamental pro-object P is isomorphic to Π as a profinite set with continuous Π-action. Thus if P' is another fundamental pro-object, we find the stabilizer of a single point p of P' has to be trivial. Since P' is connected, the action of Π is transitive; so, $\pi \mapsto \pi p$ gives an isomorphism $\Pi \cong P' \cong \Pi'$. This basically shows the equivalence as above. □

Remark 4.41 If $\mathcal{C}$ is the category of étale coverings of a connected variety S of finite type over a field k and $F(X) = F_{\overline{s}}(X) = \mathrm{Hom}_{SCH_{/S}}(\overline{s}, X)$ for a geometric point $\overline{s}$ of $S(\Omega)$, we write the fundamental group for $(\mathcal{C}, F)$ as $\pi_1(S, \overline{s})$. If one changes a point $\overline{s}$ to another geometric point $\overline{t}$, $F_{\overline{t}}$ is still a fundamental functor (see [SGA] V.7); so, we have an isomorphism $\sigma : F_{\overline{t}} \cong F_{\overline{s}}$ which induces an isomorphism again denoted by σ of the fundamental pro-objects $P_{\overline{t}} \cong P_{\overline{s}}$ associated with $F_{\overline{t}}$ and $F_{\overline{s}}$, respectively. Then $\pi_1(S, \overline{t}) \cong \pi_1(S, \overline{s})$ by $\tau \mapsto \sigma\tau\sigma^{-1}$. We call an isomorphism $\sigma : F_{\overline{t}} \cong F_{\overline{s}}$ a *path* from $\overline{s}$ to $\overline{t}$.

Remark 4.42 Suppose that S is reduced and irreducible. If we choose t to be the generic point η of S, a connected étale Galois covering $X \to S$ gives an irreducible Galois covering of S; so, it is determined by its function field $k(X)$ as a Galois extension of $k(\eta) = k(S)$. Thus $\pi_1(S, \overline{\eta})$ is canonically isomorphic to $\mathrm{Gal}(k(P)/k(S))$ for a fundamental pro-object P in $ET_{/S}$.

4.4.2 Algebraic Fundamental Groups

Let X be a locally Noetherian connected scheme. We fix a geometric point $\overline{x}$ in $X(\Omega)$. Then the category $ET_{/X}$ of étale faithfully flat coverings of X and the fundamental functor $F_{\overline{x}} : ET_{/X} \to SETS$ given by $F_{\overline{x}}(Y) = \mathrm{Hom}_{SCH_{/X}}(\overline{x}, Y)$ supply us with a Galois category and give the profinite group $\pi_1(X, \overline{x})$ so that $F_{\overline{x}}$ gives the identification of $ET_{/X}$ with $\mathcal{C}(\pi_1(X, \overline{x}))$. The group $\pi_1(X, \overline{x})$ is called the algebraic fundamental group of X with base point $\overline{x}$. We briefly expose functorial properties of the algebraic fundamental group.

When X is a smooth variety over $\mathbb{C}$, we can think of the topological fundamental group $\pi_1^{top}(X, x)$ for $x \in X(\mathbb{C})$ defined as the homotopy classes of paths starting from x ending up at x. Then taking the universal cover $p : U \to X$, any étale covering Y of X is canonically isomorphic to U/Γ for a subgroup $\Gamma = \pi_1(Y, y)$ (for a point $y \in Y(\mathbb{C})$ over x) of finite index in $\pi_1^{top}(X, x)$. Then $F_x(Y) \cong \pi_1^{top}(X, x)/\Gamma$ and thus we get

$$\pi_1(X, x) \cong \varprojlim_{\Gamma} \pi_1^{top}(X, x)/\Gamma = \widehat{\pi}_1^{top}(X, x), \tag{4.96}$$

where Γ runs over all normal subgroups of finite index in $\pi_1^{top}(X, x)$; in other words, the algebraic fundamental group $\pi_1(X, x)$ is the profinite completion of the topological fundamental group $\pi_1^{top}(X, x)$. If we pick another point $x' \in X(\mathbb{C})$ and choose a path σ from x to x', we find $\pi_1^{top}(X, x) \cong \pi_1^{top}(X, x')$ by $\tau \mapsto \sigma\tau\sigma^{-1}$, and σ induces a unique isomorphism of functors $F_x \cong F_{x'}$.

If we have a morphism $f : X \to Y$, the pullback $f^*T = X \times_Y T$ for an étale covering $T_{/Y}$ is an étale covering of X, because a pullback of a faithful flat morphism is faithfully flat (see [ECH] Corollary I.3.6). If X and Y are connected, obviously we have $F_{\overline{x}} \circ f^* \cong F_{f(\overline{x})}$ for any geometric point $\overline{x} \in X(\Omega)$, because an étale covering of a connected scheme X is determined by its fiber at any given geometric fiber ([ECH] I.3.12 and [SGA] I.5.5). Then f^* is an exact functor from $ET_{/Y}$ into $ET_{/X}$.

Suppose now that X and Y are connected locally Noetherian. Then $F_{\overline{x}} \circ f^*$ gives a covariant functor from $ET_{/Y} \to SETS$. By definition, $\pi_1(X, \overline{x}) = \mathrm{Aut}(F_{\overline{x}})$, and each $\sigma \in \pi_1(X, \overline{x})$ induces an automorphism of $F_{\overline{x}} \circ f^* \cong F_{f(\overline{x})}$. Thus $\mathrm{Aut}(F_{\overline{x}} \circ f^*) = \mathrm{Aut}(F_{f(\overline{x})}) = \pi_1(Y, f(\overline{x}))$, and we have a homomorphism of profinite groups, $\pi(f) : \pi_1(X, \overline{x}) \to \pi_1(Y, f(\overline{x}))$. In other words, the association $(X, \overline{x}) \mapsto \pi_1(X, \overline{x})$ is a covariant functor from the category of pointed schemes into the category of profinite groups.

Using this, we compare fundamental groups of the fibers of $X \xrightarrow{f} Y$ and those of X and Y. We always suppose that X and Y are connected locally Noetherian and that f is proper, surjective and separable with $f_*(\mathcal{O}_X) = \mathcal{O}_Y$ (i.e., each geometric fiber of f is connected; see [ALG] III.11.3 and [SGA] X.2). A scheme over a field k is called *separable* if for any field extension K/k, its base change to K is a reduced scheme (no nontrivial nilpotents in the structure sheaf). A morphism $X \xrightarrow{f} Y$ is called *separable* if f is flat and the

fiber $X_y = X \otimes_Y Spec(k(y))$ is separable for all points y with residue field $k(y)$ (see [SGA] X.1).

We choose a point y in the image of f in Y and a geometric point $\overline{y}$ over y so that its field $k(\overline{y})$ is an algebraic closure of $k(y)$. Take a point x in X with $f(x) = y$ and a geometric point $\overline{x}$ over x so that $k(\overline{x})$ is an algebraic closure of $k(x)$ and $f(\overline{x}) = \overline{y}$. Then we take a geometric point $\overline{a} \in \overline{X}_y = X \otimes_{k(y)} k(\overline{y})$. We write $i : \overline{X}_y \hookrightarrow X$ for the inclusion morphism. By the functoriality of π_1, we have the following sequence,

$$\pi_1(\overline{X}_y, \overline{a}) \xrightarrow{i_*} \pi_1(X, \overline{x}) \xrightarrow{f_*} \pi_1(Y, \overline{y}) \to 1, \tag{4.97}$$

which is induced by the sequence of functors $ET_{/Y} \xrightarrow{f^*} ET_{/X} \xrightarrow{i^*} ET_{/\overline{X}_y}$.

Theorem 4.43 *Suppose that X and Y are connected locally Noetherian and that f is proper, surjective, and separable with $f_*(\mathcal{O}_X) = \mathcal{O}_Y$. Then the sequence* (4.97) *is an exact sequence of profinite groups.*

Here is a sketch of the proof (see [SGA] X.1.3 for more details).

Proof. If $Y' \xrightarrow{g} Y$ is a connected étale finite covering, then f^*Y'/X is a finite étale covering and has to be connected because $f_*(\mathcal{O}_X) = \mathcal{O}_Y$. This shows that $F_{\overline{x}}(f^*Y) = F_{\overline{y}}(Y)$, and hence f^* induces a surjection of the universal pro-object of $ET_{/X}$ onto that of $ET_{/Y}$. Since the universal pro-object is in bijection with the fundamental group, we find that $\pi_1(X, \overline{x}) \xrightarrow{f_*} \pi_1(Y, \overline{y})$ is a surjection.

Since the fiber $g^{-1}(\overline{y})$ of Y' over $\overline{y}$ is a finite set of closed points, $i^* \circ f^*(Y')$ is a disjoint union of $\overline{X}_y$ indexed by points of $g^{-1}(\overline{y})$. Since $\overline{X}_y$ is the final object of $ET_{/\overline{X}_y}$, we find that $\mathrm{Im}(i_*) \subset \mathrm{Ker}(f_*)$.

To prove the reverse inclusion, $\mathrm{Im}(i_*) \supset \mathrm{Ker}(f_*)$, we need to show that for any finite étale Galois covering $X' \xrightarrow{h} X$, we have an étale Galois covering Y'/Y so that $X' \cong Y' \times_Y X$ if and only if $\overline{X}'_y/\overline{X}_y$ has a section (i.e., a disjoint union of copies of $\overline{X}_y$). The only nontrivial point is the "if" part. By the properness of f and the finiteness of h, $\mathcal{A} = (f \circ h)_*(\mathcal{O}_{X'})$ is a sheaf of $\mathcal{O}_Y$-algebras finite over $\mathcal{O}_Y$ ([EGA] III.3.2.1). Define $Y' = Spec(\mathcal{A})$ (cf. [GME] 5.4). By the separability of f, Y'/Y is an étale covering (see [EGA] III.7.8.10 and [SGA] X.2). Then the assertion is plain. □

By Hensel's lemma, if A is a Henselian local ring with residue field κ, the category of finite étale algebras over κ is equivalent to that over A by taking fibers over κ (see [ECH] I.4.2 and I.4.4). We have the following generalization of this fact to proper smooth schemes by A. Grothendieck and M. Artin:

Theorem 4.44 *Let A be a henselian local ring and $X_{/A}$ be a smooth proper scheme with special fiber $X_{0/\kappa}$. Then we have $\pi_1(X, \overline{x}) \cong \pi_1(X_0, \overline{x}_0)$ canonically, where x_0 is a geometric point in the special fiber, and $\overline{x}$ is the geometric point of X whose schematic closure contains the image of $\overline{x}_0$.*

See [ECH] Remark I.4.6 for an indication of references containing the proof.

4.4.3 Group-Theoretic Results

We gather here group-theoretic results we have used in the proof of the global reciprocity law. Here we use the notation introduced in the proof of Theorem 4.14. In particular, Σ is a proper subset of the set of rational primes. In the proof, we have used the following facts for $G_1 = \mathrm{Res}_{F/\mathbb{Q}}SL(2)$ and $PG = \mathrm{Res}_{F/\mathbb{Q}}(PGL(2))$.

(S) *For open compact subgroups $S, {}^{\sigma}S \subset G_1(\mathbb{A}^{(\Sigma\infty)})$, if $\sigma : S \cong {}^{\sigma}S$ is an isomorphism of groups, replacing S by an open subgroup, σ is induced by $g(\sigma) \in \mathrm{Aut}(F_{\mathbb{A}}^{(\Sigma\infty)}/\mathbb{A}^{(\Sigma\infty)}) \ltimes PG(\mathbb{A}^{(\Sigma\infty)})$: ${}^{\sigma}s = g(\sigma)sg(\sigma)^{-1}$.*

We may modify σ by $g \in \mathrm{Aut}(F_{\mathbb{A}}^{(\Sigma\infty)}/\mathbb{A}^{(\Sigma\infty)}) \ltimes PG(\mathbb{A}^{(\Sigma\infty)})$ so that its projection σ_F to $\mathrm{Aut}(F_{\mathbb{A}}^{(\Sigma\infty)}/\mathbb{A}^{(\Sigma\infty)})$ is the identity map. Then this assertion (S) follows from the following three assertions for σ with $\sigma_F = 1$.

(S1) For open subgroups S_ℓ and ${}^{\sigma}S_\ell$ of $G_1(\mathbb{Q}_\ell)$ (for every prime ℓ), an isomorphism $\sigma_\ell : S_\ell \cong {}^{\sigma}S_\ell$, after replacing S_ℓ by an open subgroup of S_ℓ, is induced by the conjugation $s \mapsto g_\ell(\sigma)sg_\ell(\sigma)^{-1}$ for $g_\ell(\sigma) \in G(\mathbb{Q}_\ell)$.

(S2) For a sufficiently large rational prime ℓ, we have

$$\mathrm{Aut}(G_1(\mathbb{Z}_\ell)) = \mathrm{Aut}(F_\ell/\mathbb{Q}_\ell) \ltimes PG(\mathbb{Z}_\ell)$$
$$\text{and } \mathrm{Aut}(G_1(\mathbb{Q}_\ell)) = \mathrm{Aut}(F_\ell/\mathbb{Q}_\ell) \ltimes PG(\mathbb{Q}_\ell),$$

where PG is the quotient of G modulo its center.

(S3) For sufficiently large distinct rational primes p and ℓ, any group homomorphism $\phi : G_1(\mathbb{Z}_p) \to G_1(\mathbb{Z}_\ell)$ is trivial (i.e., $\mathrm{Ker}(\phi) = G_1(\mathbb{Z}_p)$).

The assertion (S1) follows directly from a result of Lazard on ℓ-adic Lie groups (see [GAN] IV.3.2), because automorphisms of the Lie algebra of S (hence of ${}^{\sigma}S$) are all inner up to automorphisms of the field (in the cases we need). In the case of classical groups (including $SL(2)$), this can be proven easily (see the argument after Proposition 4.47).

Since the assertions (S2–3) can be proven for more general linear algebraic groups G_1 smooth over the localization $\mathbb{Z}_{(p\ell)} \subset \mathbb{Q}$ at two primes p and ℓ, we introduce some terminology in order to prove them for more general groups including $SL(2)$. For a semi-simple group G_1 over a field k, we take a maximal split torus $T^{sp} \subset G_1$ defined over k (so T^{sp} is an algebraic subgroup of G_1 maximal among those isomorphic to $\mathbb{G}_m^s$), and take a maximal torus T containing T^{sp}. The group G_1 is *split* if $T^{sp} = T$ and is *quasi-split* if the derived group of the centralizer $Z(T^{sp})$ of T^{sp} in G_1 is trivial (so, $Z(T^{sp}) = T$; see [Tt]). The only thing we need to prove (S1–3) is the fact that the structure of quasi-split groups is determined by its root system and a small amount of additional data (see [Tt]); particularly, we need the fact that the group is generated (integrally) by unipotent elements.

For a finite set of rational primes Σ', we write $\mathbb{Z}^{(\Sigma')} \subset \mathbb{Q}$ for the localization outside Σ'; so, $\mathbb{Z}^{(\Sigma')} = \mathbb{Z}\left[\frac{1}{p} | p \in \Sigma'\right]$. We suppose that $G_{1/\mathbb{Z}^{(\Sigma')}}$ is a (smooth)

semi-simple classical group; so, $G_{1/\mathbb{Q}}$ and $G_{1/\mathbb{F}_p}$ for all $p \notin \Sigma'$ are semi-simple. Since the conditions (S2–3) concern only large primes, we may assume

(QS) $G_{1/\mathbb{Z}_p}$ is smooth quasi-split for all primes p outside Σ' of rational primes (i.e., all fibers: $G_{1/\mathbb{F}_p}$ and $G_{1/\mathbb{Q}_p}$ are quasi-split for all $p \notin \Sigma'$).

We now prove the assertion (S3). Let $\phi : G_1(\mathbb{Z}_p) \to G_1(\mathbb{Z}_\ell)$ be a homomorphism. Since $G_1(\mathbb{Z}_p)$ is quasi split, $G_1(\mathbb{Z}_p)$ is generated by unipotent elements (see Lemma 4.46), and its unipotent radical U is generated by an additive subgroup U_α corresponding to a simple root α.

If $G_1 = SL(n)_{/\mathbb{Z}}$, for example, we may assume that U_α is made of $\mathrm{diag}[1_j, \left(\begin{smallmatrix} 1 & u \\ 0 & 1 \end{smallmatrix}\right), 1_{n-j-2}]$ with $u \in \mathbb{Z}_p$ for an index j (with $1 \le j \le n$).

In general, U and U_α are p-profinite. We consider the normalizer $N(U_\alpha)$ and the centralizer $Z(U_\alpha)$ of U_α in G_1. Then by conjugation, $N(U_\alpha)/Z(U_\alpha)$ acts on U_α. Since ϕ is a group homomorphism, the quotient $N(U_\alpha)/Z(U_\alpha)$ keeps acting on the image $\phi(U_\alpha)$ in $G_1(\mathbb{Z}_\ell)$ through conjugation by elements in $\phi(N(U_\alpha))$. If $p \neq \ell$, every element of $\phi(U_\alpha)$ is semi-simple (because unipotent radical of $G_1(\mathbb{Z}_\ell)$ is ℓ-profinite). Thus the centralizer (resp. the normalizer) of $\phi(U_\ell)$ is given by $Z(\mathbb{Z}_\ell)$ (resp. $N(\mathbb{Z}_\ell)$) for a reductive subgroup Z (resp. N) of G_1. Then $N(\mathbb{Z}_\ell)/Z(\mathbb{Z}_\ell)$ is a finite subgroup of the Weyl group W_1 of $G_{1/\mathbb{C}}$ which is independent of ℓ.

For example, if $G_1 = SL(n)_{/\mathbb{Z}}$, and if $\phi(U_\alpha)$ is made of diagonal matrices $\mathrm{diag}[\zeta_1 1_{m_1}, \zeta_2 1_{m_2}, \dots, \zeta_r 1_{m_r}]$ for generically distinct ζ_j, Z is given by the subgroup

$$SL(n) \cap (GL(m_1) \times GL(m_2) \times \cdots \times GL(m_r)),$$

where $GL(m_1) \times GL(m_2) \times \cdots \times GL(m_r)$ is embedded in $GL(n)$ diagonally. The quotient N/Z in this case is isomorphic to the subgroup of permutation matrices preserving Z.

If $\phi(U_\alpha)$ is nontrivial, the image of $N(U_\alpha)/Z(U_\alpha)$ in $\mathrm{Aut}(\phi(U_\alpha))$ grows at least on the order of p as p grows. In the above example of $G_1 = SL(n)_{/\mathbb{Z}}$, if $\phi(U_\alpha) \cong \mathbb{Z}/p^m\mathbb{Z}$ for $m > 0$, all elements in $\mathrm{Aut}(\phi(U_\alpha)) \cong (\mathbb{Z}/p^m\mathbb{Z})^\times$ come from $N(U_\alpha)/Z(U_\alpha) \cong \mathbb{Z}_p^\times$. This is impossible if $p \gg |W_1|$. Thus $\phi(U_\alpha) = 1$. Since $G_1(\mathbb{Z}_p)$ for large enough p is generated by U_α for all simple roots α, ϕ has to be trivial for p large enough. □

Now we prepare several lemmas and propositions in order to prove (S1–2) for $SL(2)$. Let K be a number field, and consider $V = K^2$. We consider the $\mathbb{A}$-linear endomorphism algebra $\mathrm{End}_{\mathbb{A}}(V_{\mathbb{A}})$. Let $Z(\mathbb{A})$ be the center of the $K_{\mathbb{A}}$-linear endomorphism algebra $\mathrm{End}_{K_{\mathbb{A}}}(V_{\mathbb{A}}) \subset \mathrm{End}_{\mathbb{A}}(V_{\mathbb{A}})$. Thus $K_{\mathbb{A}} \cong Z(\mathbb{A})$ by sending $f \in K_{\mathbb{A}}$ to the scalar multiplication by f.

Lemma 4.45 *The centralizer in* $\mathrm{End}_{\mathbb{A}}(V_{\mathbb{A}})$ *of* $U_{\mathbb{A}} = \left\{\left(\begin{smallmatrix} 1 & b \\ 0 & 1 \end{smallmatrix}\right) \middle| b \in K_{\mathbb{A}}\right\}$ *and an element* $\left(\begin{smallmatrix} \gamma & 0 \\ 0 & 1 \end{smallmatrix}\right) \in M_2(K)$ *with* $1 \neq \gamma \in K^\times$ *is given by* $Z(\mathbb{A})$.

Proof. By a regular representation $\rho : K_{\mathbb{A}} \hookrightarrow \mathrm{End}_{\mathbb{A}}(K_{\mathbb{A}})$, we can embed $\iota : \mathrm{End}_{K_{\mathbb{A}}}(V_{\mathbb{A}}) \hookrightarrow \mathrm{End}_{\mathbb{A}}(V_{\mathbb{A}})$. Identifying $\mathrm{End}_{\mathbb{A}}(V_{\mathbb{A}}) = M_{2d}(\mathbb{A})$ and $\mathrm{End}_{K_{\mathbb{A}}}(V_{\mathbb{A}}) =$

$M_2(K_{\mathbb{A}})$, we have $\iota \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} \rho(a) & \rho(b) \\ \rho(c) & \rho(d) \end{smallmatrix}\right)$. Let $N(A) = \left\{ \left(\begin{smallmatrix} 0 & b \\ 0 & 0 \end{smallmatrix}\right) \middle| b \in A \right\}$ for a ring A. We consider the centralizer

$$Z_1 = \{u \in \mathrm{End}_{\mathbb{A}}(V_{\mathbb{A}}) | zn = nz \quad \text{for all } n \in N(\mathbb{A})\}.$$

If $u = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in Z_1$, by computation, we find $\rho(n)d = a \cdot \rho(n)$ and $c \cdot \rho(n) = 0$ for all $n \in K_{\mathbb{A}}$; so, by taking $n = 1$, we have $a = d$ and $c = 0$. Since we can choose $\rho(n)$ with d distinct eigenvalues, we find that $a \in \rho(K_{\mathbb{A}})$. Thus we find

$$Z_1 = \left\{ \left(\begin{smallmatrix} \rho(a) & b \\ 0 & \rho(a) \end{smallmatrix}\right) \middle| a \in K_{\mathbb{A}} \quad \text{and} \quad b \in \mathrm{End}_{\mathbb{A}}(K_{\mathbb{A}}) \right\}.$$

We then consider $Z = \left\{ z \in Z_1 \middle| z \left(\begin{smallmatrix} \rho(\gamma) & 0 \\ 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} \rho(\gamma) & 0 \\ 0 & 1 \end{smallmatrix}\right) z \quad \text{for } 1 \neq \gamma \in K^\times \right\}$. Then by computation, we find $\rho(\gamma)b = b$; so, $b = 0$. This shows that Z is the center $Z(\mathbb{A})$ of $\mathrm{End}_{K_{\mathbb{A}}}(V_{\mathbb{A}})$. □

Lemma 4.46 *Let K be a field. Then $SL_2(K)$ and $PSL_2(K)$ are generated by unipotent elements. If $|K| > 2$, the commutator $[GL_2(K), GL_2(K)]$ is equal to $SL_2(K)$, and $[PGL_2(K), PGL_2(K)] = PSL_2(K)$.*

Proof. Let H be a subgroup generated by all unipotent elements in $SL_2(K)$ and $\overline{H}$ be its image in $PSL_2(K)$. If u is a nontrivial unipotent element, we find an eigenvector $v \neq 0$ of $V = K^2$ fixed by u. Taking a base (w, v) of V over K, we find $u(w, v) = (aw + bv, v) = \left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right)$. Taking the determinant, we find $a = 1$; so, u is a conjugate of $\left(\begin{smallmatrix} 1 & b \\ 0 & 1 \end{smallmatrix}\right)$. This shows that $g\overline{H}g^{-1} = \overline{H}$. Thus $\overline{H}$ is a normal subgroup of $PSL_2(K)$. Since the upper triangular unipotent group acts by translation on the affine space $\mathbf{A}^1(K) = K$, it acts transitively on $\mathbf{A}^1(K)$. The opposite unipotent subgroup ${}^tU$ acts transitively on $\mathbf{P}^1(K) - \{0\} \cong \mathbf{A}^1(K)$, and H acts transitively on $\mathbf{P}^1(K)$. Thus for any $g \in PSL_2(K)$, we find $h \in \overline{H}$ such that $h(\infty) = g(\infty)$; so, $h^{-1}g \in B$, where B is the subgroup of upper triangular matrices. Thus if $\overline{T} \subset \overline{H}$ for the diagonal torus $\overline{T}$, we find $\overline{T} \cdot \overline{H} = PSL_2(K)$.

Since $\left(\begin{smallmatrix} 1 & 0 \\ x & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & x^{-1} \\ 0 & 1 \end{smallmatrix}\right) \left(\begin{smallmatrix} 0 & -x^{-1} \\ x & 0 \end{smallmatrix}\right) \left(\begin{smallmatrix} 1 & x^{-1} \\ 0 & 1 \end{smallmatrix}\right)$, we have $j(x) = \left(\begin{smallmatrix} 0 & -x^{-1} \\ x & 0 \end{smallmatrix}\right) \in H$. Since $j(-x)j(1) = \left(\begin{smallmatrix} x^{-1} & 0 \\ 0 & x \end{smallmatrix}\right)$, we find that $T \subset H$. Thus $SL_2(K)$ and $PSL_2(K)$ are always generated by unipotent elements.

Write $[g, h] = ghg^{-1}h^{-1}$. Then we have $\left[\left(\begin{smallmatrix} a & 0 \\ 0 & b \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right)\right] = \left(\begin{smallmatrix} 1 & (a^{-1}b-1)x \\ 0 & 1 \end{smallmatrix}\right)$. Thus if $|K| > 2$, we can choose $a, b \in K^\times$ with $a \neq b$. Thus $[GL_2(K), GL_2(K)]$ contains all unipotent elements; so, $[GL_2(K), GL_2(K)] = SL_2(K)$. By the same argument, we find $[PGL_2(K), PGL_2(K)] = PSL_2(K)$. □

Fix a continuous section $s : K^\times \to GL_2(K)$ of the determinant map satisfying $x = \det(s(x))$. Let $\mathrm{Aut}_{\det}(GL_2(K))$ be the automorphism group of the group $GL_2(K)$ preserving the section s up to inner and field automorphisms, that is, $\sigma \in \mathrm{Aut}_{\det}(GL_2(K))$ satisfies $\sigma(s(a)) = h \cdot \tau(s(a))h^{-1}$ for a field automorphism $\tau \in \mathrm{Aut}(K)$ and $h \in GL_2(K)$. More generally, for a subgroup H containing $SL_2(K)$, we define $\mathrm{Aut}_{\det}(H)$ for the automorphism

group of H preserving the section s of $\det : H \to K^\times$ up to field and inner automorphisms of K. The inner automorphisms of $GL_2(K)$ give rise to an inclusion $PGL_2(K) \subset \mathrm{Aut}_{\det}(GL_2(K))$. A field automorphism $\sigma \in \mathrm{Aut}(K)$ can be considered to be a group automorphism of $GL_2(K)$ by its action entry by entry on matrices. Here is (S2) for $SL_2(K)$.

Proposition 4.47 *Let K be a non-Archimedean local field or a global field (i.e., a number field). Then the automorphism groups of the groups $SL_2(K)$, $PSL_2(K)$, and $PGL_2(K)$ are given by $\mathrm{Aut}(K) \ltimes PGL_2(K)$. If H is a subgroup of $GL_2(K)$ containing $SL_2(K)$, $\mathrm{Aut}_{\det}(H)$ is given by*

$$\left\{(\tau, g) \in \mathrm{Aut}(K) \ltimes PGL_2(K) \middle| \tau(\det(H)) = \det(H)\right\}.$$

Proof. We first study $\mathrm{Aut}(PGL_2(K))$. We have an exact sequence,

$$1 \to PGL_2(K) \xrightarrow{i} \mathrm{Aut}(PGL_2(K)) \to \mathrm{Out}(PGL_2(K)) \to 1,$$

where $i(x)(g) = xgx^{-1}$. We write B (resp. U) for the upper triangular Borel subgroup (resp. the upper triangular unipotent subgroup) of $GL_2(K)$. Their image in $PGL_2(K)$ is denoted by $\overline{B}$ and $\overline{U}$. Let U' be a subgroup of $GL_n(K)$ isomorphic to the additive group K; so, we have $a : K \cong U'$. Consider the image $a(1)$ of $1 \in K$ in U'. Replacing K by a finite extension containing an eigenvalue α of $a(1)$, let $V_\alpha \subset V$ be the eigenspace of $a(1)$ with eigenvalue α. Then $a(\frac{1}{m})$ acts on V_α and $a(\frac{1}{m})^m = a(1) = \alpha \in \mathrm{End}(V_\alpha)$. Thus we have an algebra homomorphism $K[x]/(x^m - \alpha) \to \mathrm{End}_K(V_\alpha)$ for all $0 < m \in \mathbb{Z}$. If K is a p-adic local field, $\bigcap_m (K^\times)^{m!} = \{1\}$. By using this, we find $\bigcap_m (K^\times)^m = \{1\}$ for a global field K. Thus if K is a non-Archimedean local field or a global field, we find that α has to be 1. Thus U' is made up of commuting unipotent elements; so, by conjugation, we can bring U' onto U.

Since $\sigma(\overline{U})$ for $\sigma \in \mathrm{Aut}(PGL_2(K))$ is isomorphic to K in $PGL_2(K)$, $\sigma(\overline{U})$ is again a unipotent subgroup of $PGL_2(K)$. Since $\overline{B}$ is the normalizer of $\overline{U}$, again $\sigma(\overline{B})$ is the normalizer of $\sigma(\overline{U})$; so, $\sigma(\overline{B})$ is a Borel subgroup. We find $g \in GL_2(K)$ such that $\sigma(\overline{B}) = g\overline{B}g^{-1}$. Thus we may assume that σ fixes $\overline{B}$. Applying the same argument to $\overline{U}$, we may assume that σ fixes $\overline{U}$. Let $\sigma\left(\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right) = \left(\begin{smallmatrix}1 & \alpha\\ 0 & 1\end{smallmatrix}\right)$. Then changing σ by $i\left(\begin{smallmatrix}\alpha^{-1} & 0\\ 0 & 1\end{smallmatrix}\right) \circ \sigma$, we may assume that $\sigma\left(\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right) = \left(\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right)$. Then σ induces an automorphism $\overline{\sigma}$ of $T = \overline{B}/\overline{U}$. Fix an isomorphism $t : K^\times \cong T$ and $u : K \cong \overline{U}$. Since σ sends the center to the center, we find $t(a)u(b)t(a)^{-1} = u(ab)$. Then applying σ to the above formula, we get $\overline{\sigma}(a)\sigma(b) = \sigma(ab)$ for $a \in K^\times$ and $b \in K$. Taking $b = 1$, we find $\overline{\sigma}(a) = \sigma(a)$. Thus $\sigma(ab) = \sigma(a)\sigma(b)$, and $\sigma : K \to K$ is a field automorphism. Modifying σ by an element of $\mathrm{Aut}(K)$, we may assume that σ induces the identity map on B. Thus σ induces an automorphism of $\mathbf{P}^1(K) = GL_2(K)/B$ fixing ∞ and commuting with all affine transformations $z \mapsto az+b$ of $\mathbf{P}^1(K)$ for $a \in K^\times$ and $b \in K$. Then σ is the identity map on $\mathbf{P}^1(K)$ and, hence, on $GL_2(K)$. We find $\mathrm{Out}(PGL_2(K)) \cong \mathrm{Aut}(K)$ and $\mathrm{Aut}(PGL_2(K)) = \mathrm{Aut}(K) \ltimes PGL_2(K)$.

We apply the same argument to $\mathrm{Aut}(PSL_2(K))$. Let $\overline{B}_1 = \overline{B} \cap PSL_2(K)$ and $\overline{T}_1 = \overline{T} \cap PSL_2(K)$. We find $g \in PSL_2(K)$ such that $\sigma(B_1) = gB_1g^{-1}$. Thus we may assume that σ leaves $\overline{B}_1$ and $\overline{U}$ stable. Further modifying σ by conjugation of an element in $GL_2(K)$, we may assume that $\sigma(u(1)) = u(1)$. Note that in this case, $t(a)u(b)t(a)^{-1} = u(a^2b)$. Then we have $\overline{\sigma}(a^2)\sigma(b) = \sigma(a^2b)$; so, $\sigma(a^2) = \overline{\sigma}(a^2) = \overline{\sigma}(a)^2$. Thus $\overline{\sigma}(a) = \pm\sigma(a)$. Thus if K is a field of characteristic 2, we find that $\sigma = \overline{\sigma}$ is a field automorphism of K.

Suppose that K has a characteristic different from 2. This implies

$$\sigma(a^2) = \overline{\sigma}(a^2) = \overline{\sigma}(a)^2 = (\pm\sigma(a))^2 = \sigma(a)^2,$$

$\sigma((a+b)^2) = \sigma(a^2 + 2ab + b^2) = \sigma(a)^2 + 2\sigma(ab) + \sigma(b)^2$ and $\sigma((a+b)^2) = \overline{\sigma}(a+b)^2 = \sigma(a+b)^2 = \sigma(a)^2 + 2\sigma(a)\sigma(b) + \sigma(b)^2$. Thus if K has characteristic different from 2, we find $\sigma(ab) = \sigma(a)\sigma(b)$; so, $\sigma \in \mathrm{Aut}(K)$. Thus for a character $\varepsilon : K^\times/(K^\times)^2 \to \{\pm 1\}$, we have $\overline{\sigma}(x) = \varepsilon(x)\sigma(x)$; so, $\sigma(t(a)) = t(\overline{\sigma}(a)) = \pm t(\sigma(a)) = t(\sigma(a))$ in $PSL_2(K)$. This shows $\mathrm{Aut}(PSL_2(K)) = \mathrm{Aut}(K) \ltimes PGL_2(K) = \mathrm{Aut}(PGL_2(K))$.

We now study $\mathrm{Aut}(SL_2(K))$. Since physically $SL_2(K) = PSL_2(K)$ for a field of characteristic 2, we may assume that K has a characteristic different from 2. We have an exact sequence

$$1 \to \{\pm 1\} \to SL_2(K) \to PSL_2(K) \to 1.$$

Since $\mathrm{Aut}(K) \ltimes PGL_2(K) \subset \mathrm{Aut}(SL_2(K))$, each automorphism σ of $PSL_2(K)$ lifts to an automorphism of $SL_2(K)$, which in turn induces an automorphism of $\{\pm 1\}$. Since $\mathrm{Aut}(\{\pm 1\}) = \{1\}$, we find that $\mathrm{Aut}(SL_2(K)) = \mathrm{Aut}(PSL_2(K)) = \mathrm{Aut}(K) \ltimes PGL_2(K)$.

Now we look at $\mathrm{Aut}_{\det}(H)$. Since $\sigma \in \mathrm{Aut}_{\det}(H)$ preserves the section s up to field and inner automorphisms, modifying σ by such an automorphism, we may assume that σ fixes $s(K^\times)$. Then σ is determined by its restriction to $SL_2(K) \subset H$ and, hence, comes from and element in $\mathrm{Aut}(K) \ltimes PGL_2(K)$ preserving H. This shows the last assertion. □

Here is a direct proof of (S1) for $SL_2(K)$ for a finite extension K of $\mathbb{Q}_p$, although the proof in [GAN] is a highly polished and generalized version of this argument. The exponential $\exp(X) = \sum_{n\geq 0} \frac{X^n}{n!}$ and the logarithm $\log(1+X) = \sum_{n\geq 1}(-1)^{n+1}\frac{X^n}{n}$ converge p-adically on a sufficiently small neighborhood of 0 of $M_2(K)$ and 1 of $GL_2(K)$, respectively. If $\sigma : S \cong S'$ is an isomorphism of open compact subgroups of $SL_2(K)$, by shrinking S and S', we may assume that the log induces an isomorphism $\log : S \cong \log(S)$ and $S' \cong \log(S')$. Then σ induces an isomorphism of the Lie algebra $\sigma : \log(S) \cong \log(S')$, which extends to an automorphism of the Lie algebra $\mathfrak{sl}_2(K)$ of $SL_2(K)$ by scalar extension. Since σ takes the maximal abelian Lie algebra $N(K) \cong K$ of $\mathfrak{sl}_2(K)$ into its conjugate, changing σ by σ composed with a conjugation by an element of $GL_2(K)$, we may assume that σ takes $N(K)$ onto $N(K)$. Identifying $N(K)$ with K by sending each matrix in $N(K)$ to its

upper-right corner element, as in the proof of the above proposition, we have $\sigma(ab) = \sigma(a)\sigma(b)$ for $a, b \in K$. Thus by Lemma 4.45 (or its proof), we may assume that $\sigma \in \mathrm{End}(\mathfrak{sl}_2(K))$ is K-linear. Then by the Lie algebra version of Proposition 4.47, it is a conjugation by an element of $GL_2(K)$. This shows (S1) for SL_2. This argument proving (S1) can be generalized to quasi-split semi-simple classical groups more general than $SL(2)$.

Here is (S2) for $SL_2(O)$.

Proposition 4.48 *Suppose that K is a finite extension of $\mathbb{Q}_p$ with integer ring O. Then the automorphism groups of the groups $PSL_2(O)$, $SL_2(O)$, and $PGL_2(O))$ are given by $\mathrm{Aut}(O) \ltimes PGL_2(O)$. If H is a subgroup of $GL_2(O)$ containing $SL_2(O)$, $\mathrm{Aut}_{\det}(H)$ is given by*

$$\left\{(\tau, g) \in \mathrm{Aut}(K) \ltimes PGL_2(O) \middle| \tau(\det(H)) = \det(H)\right\}.$$

Proof. Let κ be the residue field of O. We study $\mathrm{Aut}(PGL_2(O))$. We have $\mathrm{Aut}(K) \ltimes PGL_2(O) \subset \mathrm{Aut}(SL_2(O))$. Let $U(O) = U \cap SL_2(O)$. For the maximal ideal $\mathfrak{m} \subset O$, we define $\Gamma(\mathfrak{m}^n) = \left\{\alpha \in SL_2(O) \middle| \alpha \equiv 1 \mod \mathfrak{m}^n\right\}$. Since $\mathbf{P}^1(O) = \mathbf{P}^1(K)$, all Borel subgroups of $SL_2(O)$ are conjugates of each other. Since $B_1(O) = SL_2(O) \cap B$ is a semi-direct product of $T_1(O)$ and $U(O)$, all unipotent subgroups are conjugates of each other. By the same argument in the case of the field, we may assume that $\sigma \in \mathrm{Aut}(PGL_2(O))$ leaves $U(O)$ stable. We have $t(a)u(b)t(a)^{-1} = u(ab)$. Then applying σ to the above formula, we get $\overline{\sigma}(a)\sigma(b) = \sigma(ab)$ for $a \in K^\times$ and $b \in K$. If $\sigma(1) \in \mathfrak{m}$, then $\overline{\sigma}(a)\sigma(1) = \sigma(ab) \in \mathfrak{m}$. However, $\{\sigma(a) | a \in O\} = O$; we find $O \subset \mathfrak{m}$, a contradiction. Thus $\sigma(1) = \alpha$ is a unit. Then $t(\alpha) \in GL_2(O)$, and modifying $i(\alpha)$, we may assume that $\sigma(1) = 1$. Proceeding in exactly the same way in the case of the field, we find that $\mathrm{Aut}(PGL_2(O)) = \mathrm{Aut}(O) \ltimes PGL_2(O)$. From this, again we obtain all the desired assertions. □

Our argument proving Propositions 4.47 and 4.48 can be easily generalized to a semi-simple algebra K of finite dimension over $\mathbb{Q}$ or $\mathbb{Q}_p$. We only state the result leaving the proof to attentive readers.

Proposition 4.49 *Assume that K is a semi-simple commutative algebra of finite dimension over $\mathbb{Q}$ or $\mathbb{Q}_p$ (or the direct sum of such). Then the automorphism groups of the groups $PGL_2(K)$, $PSL_2(K)$, and $SL_2(K)$ are given by $\mathrm{Aut}(K) \ltimes PGL_2(K)$. If K is a semi-simple $\mathbb{Q}_p$-algebra, we again have*

$$\mathrm{Aut}(PGL_2(O)) = \mathrm{Aut}(PSL_2(O)) = \mathrm{Aut}(SL_2(O)) = \mathrm{Aut}(O) \ltimes PGL_2(O),$$

where O is the integral closure of $\mathbb{Z}_p$ in K.

The group $\mathrm{Aut}_{\det}(H)$ for a subgroup of $GL_2(K)$ and $GL_2(O)$ containing $SL(2)$ has the same expression as in Propositions 4.47 and 4.48.

A similar argument proves (S2) for symplectic and unitary groups, for whose proof we refer to [H03b] Section 3. Let $J_{2n} = \left(\begin{smallmatrix} 0 & -1_n \\ 1_n & 0 \end{smallmatrix}\right)$, $I_{2n} = \left(\begin{smallmatrix} 0 & 1_n \\ 1_n & 0 \end{smallmatrix}\right)$,

and $I_{2n+1} = \begin{pmatrix} 0 & 0 & 1_n \\ 0 & 1 & 0 \\ 1_n & 0 & 0 \end{pmatrix}$. We take a finite extension field K of $\mathbb{Q}_p$ and its semisimple quadratic extension M; so, M is either a field extension or isomorphic to $K \oplus K$. We write R (resp. O) for the integral closure of $\mathbb{Z}_p$ in M (resp. K). Write c for the generator of the ring automorphism group $\mathrm{Aut}(M/K)$. Define $G(A)$ for $A = K$ and O by one of the following groups,

$$\begin{aligned} G(A) &= \left\{\alpha \in GL_n(A \otimes_O R) \middle| \alpha I_n \cdot {}^t\alpha^c = \nu(\alpha) I_n\right\}, \\ G(A) &= \left\{\alpha \in GL_{2n}(A) \middle| \alpha J_{2n} \cdot {}^t\alpha = \nu(\alpha) J_{2n}\right\}, \end{aligned} \tag{4.98}$$

where $\nu : G(A) \to A^\times$ is a group homomorphism (called the similitude map). We call $G_{/K}$ a quasi-split *unitary* group if it is defined for M/K and I_n. We write $G = GU(I_n)$ in this case. The group defined for J_{2n} is called the *symplectic* group. We write $G = GSp(2n)$ in this case. We have the determinant map $\det : G(K) \to M^\times$ for unitary groups G and write π for $\det \times \nu : G(K) \to M^\times \times K^\times$. We then define PG by G modulo its center and write G_1 for $\mathrm{Ker}(\pi)$ or $\mathrm{Ker}(\nu)$ accordingly as $G = GU(I_n)$ or $G = GSp(2n)$. Fixing a section s of $? = \pi$ or ν, we define $\mathrm{Aut}_?(G(A))$ similarly to $\mathrm{Aut}_{\det}(GL_2(A))$. We assume that $p > 2$ is unramified in $M/\mathbb{Q}_p$.

Theorem 4.50 *Under the above notation and assumption, we have*

$$\begin{aligned} \mathrm{Aut}_?(G(A)) = \mathrm{Aut}(PG(A)) = \mathrm{Aut}(G_1(A)) \\ = \begin{cases} \mathrm{Aut}(K) \ltimes PG(A) & \text{if } G = GSp(2n), \\ \mathrm{Aut}(M) \ltimes PG(A) & \text{if } G = GU(I_n), \end{cases} \end{aligned}$$

where $? = \nu$ *or* π *accordingly as* G *is symplectic or unitary.*

As a corollary, we have

Corollary 4.51 *If* H *is a subgroup of* $G(O)$ *containing* $G_1(O)$, $\mathrm{Aut}_\nu(H)$ *is given by*

$$\left\{(\tau, g) \in \mathrm{Aut}(K) \ltimes PG(O) \middle| \tau(\nu(H)) = \nu(H)\right\},$$

and $\mathrm{Aut}_\pi(H)$ *is given by*

$$\left\{(\tau, g) \in \mathrm{Aut}(M) \ltimes PG(O) \middle| \tau(\pi(H)) = \pi(H)\right\}.$$

5

Generalized Eichler–Shimura Map

Let $G_{/\mathbb{Z}_{(p)}}$ for $\mathbb{Z}_{(p)} = \mathbb{Q} \cap \mathbb{Z}_p$ be a connected reductive group split over $\mathbb{Z}_p$ (see Section 4.4.3 for the definition of split algebraic groups). We first prove semi-simplicity of the commutative Hecke algebra acting on the nearly ordinary cohomology group $H^q_{n.ord}(X(S), L)$ inside the interior cohomology $H^q_!(X(S), L)$ for a modular variety $X(S)$ associated with an arbitrary p-power level open compact subgroup S of $G(\mathbb{A}^\infty)$, where $H^q_!$ is the image of the compactly supported cohomology group inside the standard cohomology group. Here the locally constant or coherent sheaf L of $\mathbb{Q}$-vector spaces on $X(S)$ is associated with a rational representation of G twisted by a finite-order character. After dealing with topological cohomology groups, we relate as Hecke modules the topological and the coherent cohomology groups via the generalized Eichler–Shimura map (which shows the semi-simplicity of the Hecke algebra acting on the coherent cohomology and topological cohomology). Although we have assumed that G is split over $\mathbb{Z}_p$, the argument works equally well for a connected reductive G smooth quasi-split over $\mathbb{Z}_p$.

We write B (resp. N) for a fixed Borel subgroup of G (resp. the unipotent radical N of B). We have the torus $T_G = B/N$. In this chapter, the character U is often used to denote a unitary group.

5.1 Semi-Simplicity of Hecke Algebras

We describe the semi-simplicity of the Hecke algebra acting on topological cohomology groups. Since in this book we do not touch upon representation theory of p-adic reductive groups, purely representation-theoretic results are quoted from appropriate sources. A main point is to prove, assuming representation theoretic results, that the nearly p-ordinary part of the cohomology (where certain Hecke operators at the primes p dividing the level S act by p-integral automorphisms) is semi-simple as Hecke modules. Thus the part of the cohomology of level S (where the group $G(\mathbb{Q}_p)$ acts through nearly

p-ordinary representations) is semi-simple as Hecke modules (over an appropriate Hecke algebra).

5.1.1 Jacquet Modules

We summarize here known results on Jacquet modules from representation theory. For a p-adic Lie group $\mathcal{G}$ and a (right) representation V with coefficients in a field K of characteristic 0, we call V *smooth* if for each $v \in V$, we find an open subgroup $S \subset \mathcal{G}$ that leaves v invariant. If furthermore, $H^0(S, V)$ is finite-dimensional for all open subgroups S, V is called *admissible*. Let π be an admissible semi-simple representation of $G(\mathbb{Q}_p)$ on a vector space $V = V(\pi)$ over a field K of characteristic 0. Contrary to the tradition, we always suppose that V is a right $G(\mathbb{Q}_p)$-module.

We have a Haar measure du of $N(\mathbb{Q}_p)$ with $\int_{N(\mathbb{Z}_p)} du = 1$. We define $V(B) = V(B, \pi) = \left\{v - v\pi(n) \in V(\pi) \middle| v \in V(\pi),\ n \in N(\mathbb{Q}_p)\right\}$ as a $T_G(\mathbb{Q}_p)$-module, and put $V_B = V_B(\pi) = V/V(B)$, which is called the *Jacquet module*. By this definition, $V \mapsto V_B$ is a right exact functor. For each $w = v - v\pi(n) \in V(B)$, we take a sufficiently large open compact subgroup $U_w \subset N(\mathbb{Q}_p)$ containing n. Then we see that $\int_U v\pi(u)du = 0$ for every open subgroup U with $U_w \subset U \subset N(\mathbb{Q}_p)$. If $\int_U v\pi(u)du = 0$ for every sufficiently large open subgroup U of $N(\mathbb{Q}_p)$, for the stabilizer U' of v in U, we find

$$\left(\int_{U'} du\right) v = \left(\int_{U'} du\right) v - \int_U v\pi(u)du = \left(\int_{U'} du\right) \sum_{n \in U/U'} (v - v\pi(n))$$

which is in $V(B)$. Thus $V(B)$ is the collection of v with $\int_U v\pi(u)du = 0$ for every sufficiently large open subgroup U of $N(\mathbb{Q}_p)$. By this fact, the functor $V \mapsto V_B$ is left exact, and we conclude that the association is an exact functor. Later we give a canonical splitting $V^N \cong V_B \oplus V(B)^N$ as Hecke modules (due to Bernstein and Casselman), where $V^N = H^0(N(\mathbb{Z}_p), V)$.

For a smooth representation V' of $T_G(\mathbb{Q}_p)$ (regarding it as a B-module via the projection $B \twoheadrightarrow T_G$), a *smooth induction* from B of V' is defined by

$$\mathrm{Ind}_B^G(V') = \{f : G(\mathbb{Q}_p) \to V' | f\text{: smooth},\ f(gb) = g(g)b\ \forall b \in B(\mathbb{Q}_p)\}, \quad (5.1)$$

on which we let $G(\mathbb{Q}_p)$ act by $f(x)g = f(gx)$. Here the word "smooth" means that for each $f \in \mathrm{Ind}_B^G V'$, we find an open compact subgroup S such that $f(kx) = f(x)$ for all $k \in S$. Thus $\mathrm{Ind}_B^G V'$ is automatically smooth.

Since the smooth induction preserves admissibility ([BeZ] 2.3), $V = \mathrm{Ind}_B^G V'$ has composition series $\{0\} \subset V_1 \subset V_2 \subset \cdots \subset V$ stable under the action of G, and hence its semi-simplification $(\mathrm{Ind}_B^G V')^{ss} = \bigoplus_j V_{j+1}/V_j$ is well-defined as an admissible $G(\mathbb{Q}_p)$-module. The beauty of the theory of admissible representations is its purely algebraic nature; so, we do not need to assume any analytic assumptions; in particular, our representations are often

not unitary as is clear from our main result (Theorem 5.3) in this section. Hereafter all representations of G, B, and T_G are supposed to be admissible.

The following results are due to Jacquet and Bernstein–Zelevinsky and are well known (see [BeZ]),

1. (Frobenius reciprocity; [BeZ] 1.9) $\mathrm{Hom}_B(V_B, V') \cong \mathrm{Hom}_G(V, \mathrm{Ind}_B^G V')$;
2. If π is absolutely irreducible, then $\dim_K V_B \leq |\mathfrak{W}|$, where $\mathfrak{W}$ is the Weyl group of T_G in G (Bernstein–Zelevinsky; [BeZ] 2.9);
3. If π is absolutely irreducible and $V_B \neq 0$, then we have a surjective linear map $\mathrm{Ind}_B^G \lambda \twoheadrightarrow V$ of $G(\mathbb{Q}_p)$-modules for a character $\lambda : T_G(\mathbb{Q}_p) \to K^\times$ (Jacquet; [BeZ] 2.4–5);
4. $\left(\mathrm{Ind}_B^G \widetilde{\lambda}\right)^{ss} \cong \left(\mathrm{Ind}_B^G \widetilde{\lambda^w}\right)^{ss}$ for $w \in \mathfrak{W}$ (Bernstein–Zelevinsky; [BeZ] 2.9),

where ss indicates "semi-simplification", $\lambda^w(t) = \lambda(wtw^{-1})$, and $\widetilde{\lambda} = \delta_B^{1/2}\lambda$ for the module character δ_B of B: $\int_{N(\mathbb{Q}_p)} \phi(u)du = \delta_B(b) \int_{N(\mathbb{Q}_p)} \phi(b^{-1}xb)du$ (for all ϕ). We have the following corollary of the above facts, which is a key to our semi-simplicity result.

Proposition 5.1 *Suppose that π is irreducible and that $V_B[\widetilde{\lambda}] \neq 0$, where $\widetilde{\lambda} = \delta_B^{1/2}\lambda$ for the module character δ_B on B. Then π is a quotient of $\mathrm{Ind}_{B(\mathbb{Q}_p)}^{G(\mathbb{Q}_p)} \widetilde{\lambda}$. If $\widetilde{\lambda^w}(t)$ for $w \in \mathfrak{W}$ are all distinct, $V_B \subset \bigoplus_{w\in\mathfrak{W}} \widetilde{\lambda^w}$ as $T_G(\mathbb{Q}_p)$-modules.*

We say a representation π is *regularly induced* (at p) if for each irreducible component V of π, we have a nontrivial $\widetilde{\lambda}$-eigenspace $V_B[\widetilde{\lambda}] \neq 0$ for a character λ of $T_G(\mathbb{Q}_p)$ and $\widetilde{\lambda^w}(t)$ $(w \in \mathfrak{W})$ are all distinct.

Proof. Since the algebra in $\mathrm{End}_K(V_B)$ generated by the action of T_G is a finite-dimensional commutative algebra, the $\widetilde{\lambda}$-eigenspace is nontrivial if and only if the maximal $\widetilde{\lambda}$-quotient is nontrivial. Thus, we have a morphism of T_G-modules, $V_B \twoheadrightarrow V(\widetilde{\lambda})$. Since we have $(\mathrm{Ind}_B^G \widetilde{\lambda})^{ss} \cong (\mathrm{Ind}_B^G \widetilde{\lambda^w})^{ss}$, by Frobenius reciprocity, all eigenvalues $\widetilde{\lambda^w}$ can show up as a quotient of $(\mathrm{Ind}_B^G \widetilde{\lambda})_B$ whose dimension is bounded by $|\mathfrak{W}|$. Thus if all characters $\widetilde{\lambda^w}$ are distinct, we have $V_B \subset (\mathrm{Ind}_B^G \widetilde{\lambda})_B \cong \bigoplus_{w\in\mathfrak{W}} \widetilde{\lambda^w}$. Since $V \mapsto V_B$ is exact, this is enough to conclude the assertion. □

5.1.2 Double Coset Algebras

We consider the double coset algebra made of formal linear combinations of double cosets of a subgroup in a semi-group. This type of algebra is considered in [IAT] 3.1 and often called a Hecke ring. We use the terminology "double coset algebra" to avoid confusion with Hecke algebras (which is the image of the double coset algebra in the endomorphism algebra of a space of automorphic forms).

Let $D = \left\{x \in T_G(\mathbb{Q}_p) \middle| xN(\mathbb{Z}_p)x^{-1} \supset N(\mathbb{Z}_p)\right\}$ which is called the *expanding semi-group* in $T_G(\mathbb{Q}_p)$. Since $xN(\mathbb{Z}_p) \supset N(\mathbb{Z}_p)x$ $(x \in D)$, $\Delta_N = N \cdot D \cdot N$

for $N = N(\mathbb{Z}_p)$ is also a multiplicative semi-group. When $G = GL(n)_{/\mathbb{Z}_p}$, it is easy to see that D is generated by $T_G(\mathbb{Z}_p)$ and $\mathrm{diag}[p^{e_1}, p^{e_2}, \dots, p^{e_n}]$ for sequences of integers $e_1 \le e_2 \le \cdots \le e_n$. Write $B = B(\mathbb{Z}_p)$ and $N = N(\mathbb{Z}_p)$ for simplicity. Define the so-called Iwahori subgroups by

$$\begin{aligned} S_0(r) &= \left\{ u \in G(\mathbb{Z}_p) \middle| u \mod p^r \in B(\mathbb{Z}/p^r\mathbb{Z}) \right\}, \\ S_1(r) &= \left\{ u \in G(\mathbb{Z}_p) \middle| u \mod p^r \in N(\mathbb{Z}/p^r\mathbb{Z}) \right\}. \end{aligned} \tag{5.2}$$

These subgroups S have the Iwahori decomposition $S = N'T'N \cong N' \times T' \times N$ for open compact subgroups $T' \subset T_G(\mathbb{Z}_p)$ and N' in the opposite unipotent ${}^tN = {}^tN(\mathbb{Z}_p)$. Each $x \in D$ shrinks ${}^tN$: $x^tNx^{-1} \subset {}^tN$. From this, $\Delta_S = S \cdot D \cdot S$ is again a multiplicative subsemi-group of $G(\mathbb{Q}_p)$ (this statement includes $\Delta_B = B \cdot D \cdot B = \Delta_{S_0(\infty)}$). We call Δ_S the *expanding* semi-group with respect to (S, N). When $G = GL(2)$, Δ_S for $S = S_0(r)$ $(r > 0)$ is almost equal to the p-part of the semi-group $\Delta_0(p^r\mathfrak{N})$ introduced in Section 4.2.5 (strictly speaking, we have $\Delta_0(p^r\mathfrak{N})_p = \Delta_S$ modulo center).

By the Iwahori decomposition, we have, for $X = N(\mathbb{Z}_p), B(\mathbb{Z}_p)$, and an Iwahori subgroup S,

$$X\xi X = \bigsqcup_{u \in \xi^{-1}N\xi \backslash N} X\xi u = \bigsqcup_{u \in N \backslash \xi N \xi^{-1}} Xu\xi. \tag{5.3}$$

By this fact, the double coset algebras generated additively over $\mathbb{Z}$ by double cosets of X in Δ_X have the following homomorphic relations as algebras:

$$\mathcal{R} = R(N, \Delta_N) \twoheadrightarrow R(B, \Delta_B) \cong R(S, \Delta_S) =: \mathcal{R}_B$$

via $N\xi N \mapsto B\xi B \mapsto S\xi S$ for $\xi \in \Delta_N$. These algebras are commutative: $T(\xi)T(\eta) = T(\xi\eta)$ for $T(\xi) = N\xi N$ and $\xi, \eta \in D$. See [IAT] Chapter 3 and [MFG] 3.1.6 for this type of results dealing with $GL(2)$. Here we give a proof of these facts for $G = GL(n)$. The same argument works well for classical Chevalley groups quasi-split over $\mathbb{Z}_p$ replacing α_j below by the image of p under simple coroot α positive with respect to D (i.e., $\alpha(p) \in D$).

Lemma 5.2 *Let the notation and the assumption be as above. Assume that $G = GL(n)_{/\mathbb{Z}_p}$. The algebras $R(S, \Delta_S)$ for Iwahori subgroups S are commutative, and if $S \supset B(\mathbb{Z}_p)$, they are all isomorphic to the polynomial ring $\mathbb{Z}[t_1, \dots, t_n, t_n^{-1}]$ (with t_n inverted) for $t_j = S\alpha_j S$ with $\alpha_j = \mathrm{diag}[1_{n-j}, p1_j]$. If $S \supset N(\mathbb{Z}_p)$, we have $R(S, \Delta_S) \cong \mathbb{Z}[T_S][t_1, \dots, t_n, t_n^{-1}]$ for the quotient group $T_S = T_G(\mathbb{Z}_p)S/S$ of $T_G(\mathbb{Z}_p)$, where $\mathbb{Z}[T_S]$ is the group algebra of T_S.*

Proof. For simplicity, we write B for $B(\mathbb{Z}_p)$. For $\xi \in D$, we consider the double coset $B\xi B$. Decompose $B = \bigsqcup_{\eta \in \Xi(\xi)} (\xi^{-1}B\xi \cap B)\eta$. Multiplying by $\xi^{-1}B\xi$ from the left, we get $\xi^{-1}B\xi B = \bigsqcup_{\eta \in \Xi(\xi)} \xi^{-1}B\xi\eta \Leftrightarrow B\xi B = \bigsqcup_{\eta \in \Xi(\xi)} B\xi\eta$. If $\xi = \mathrm{diag}[p^{a_1}, \dots, p^{a_n}]$, we have

$$\xi^{-1}B\xi \cap B = \left\{ (\gamma_{ij}) \in B \middle| \gamma_{ij} \in p^{a_j - a_i}\mathbb{Z}_p \text{ for all } j > i \right\}.$$

We may choose the subset $\Xi(\xi)$ inside $N(\mathbb{Z}_p)$ so that

$$\Xi(\xi) \ni (\eta_{ij}) \mapsto \left(\eta_{ij} \mod p^{a_j - a_i}\mathbb{Z}_p\right)_{j>i} \in \bigoplus_{j>i} \mathbb{Z}/p^{a_j-a_i}\mathbb{Z}$$

is a bijection. Then we have $B\xi B = \bigsqcup_{\eta\in\Xi(\xi)} B\xi\eta$ and a formula of the index: $(B : \xi^{-1}B\xi \cap B) = p^{[\xi]}$ with $[\xi] = \sum_{j>i}(a_j - a_i)$. Writing $\deg(B\xi B)$ for the number of left cosets of B in $B\xi B$, we find $\deg(B\xi\zeta B) = \deg(B\xi B)\deg(B\zeta B)$, because $[\xi\zeta] = [\xi] + [\zeta]$ for $\xi, \zeta \in D$. Since $B\xi B\zeta B \supset B\xi\zeta B$, if we can show that $\deg(B\xi B\zeta B) = \deg(B\xi\zeta B)$, we get $B\xi B\zeta B = B\xi\zeta B$ and $(B\xi B)\cdot(B\zeta B) = B\xi\zeta B$ in the double coset ring $R(B, \Delta_B)$, which in particular shows the commutativity of $R(B, \Delta_B)$. To see $\deg(B\xi B\zeta B) = \deg(B\xi\zeta B)$, we note $B\xi B\zeta B = \bigcup_{\eta\in\Xi(\xi)}\bigcup_{\eta'\in\Xi(\zeta)} B\xi\eta\zeta\eta'$. This implies

$$\deg(B\xi B\zeta B) \le \deg(B\xi B)\deg(B\zeta B) = \deg(B\xi\zeta B),$$

and hence we get the identity $\deg(B\xi B\zeta B) = \deg(B\xi\zeta B)$. Since the α_js give independent generators of $D/T_G(\mathbb{Z}_p)$, the monoid algebra $\mathbb{Z}[D/T_G(\mathbb{Z}_p)]$ is isomorphic to a polynomial ring with n variables $\mathbb{Z}[\alpha_1, \dots, \alpha_n, \alpha_n^{-1}]$ with α_n inverted (because $\alpha_n \in Z(\mathbb{Q}_p)$ is invertible in $Z(\mathbb{Q}_p) \subset D$). The association $\alpha_j \mapsto T(\alpha_j) = B\alpha_j B$ therefore induces a surjective algebra homomorphism $\mathbb{Z}[D/T_G(\mathbb{Z}_p)] \to R(B, \Delta_B)$, which can be easily seen to be an isomorphism. Replacing $D/T_G(\mathbb{Z}_p)$ by $D/(T_G(\mathbb{Z}_p) \cap S) = T_S \times (D/T_G(\mathbb{Z}_p))$ in the above argument, the same proof works well for any S with $S_0(r) \supset S \supset S_1(r)$ and yields $R(S, \Delta_S) \cong \mathbb{Z}[T_S][D/T_G(\mathbb{Z}_p)] \cong \mathbb{Z}[T_S][t_1, \dots, t_n, t_n^{-1}]$, where $\mathbb{Z}[T_S]$ is embedded into $R(S, \Delta_S)$ by sending $t \in T_S$ to StS. □

We let $\mathcal{R}$ act on $v \in V^N = H^0(N(\mathbb{Z}_p), V)$ by

$$v|T(\xi) = v|[N\xi N] = \sum_{u\in\xi^{-1}N\xi\backslash N} v\pi(\xi u) = \int_{\xi N\xi^{-1}} v\pi(u)\pi(\xi)du, \tag{5.4}$$

and similarly for $\overline{v} \in V_B$ in place of $v \in V^N$; then the projection: $V^N \to V_B$ is $\mathcal{R}$-linear. To see the last identity of (5.4), it is sufficient to recall that we have normalized the measure du so that $\int_N du = 1$.

Let Σ be the set of maximal (proper) parabolic subgroups $P \supset B$. Decomposing $P = M_P N_P$ for a reductive part $M_P \supset T_G$ (called a Levi factor; see [RAG] II.1.8) and the unipotent radical N_P, we can identify the set Σ with the following set of cocharacters,

$$\left\{\alpha_P : \mathbb{G}_m \to G \middle| \alpha_P(p) \text{ generates } Z_P(\mathbb{Q}_p) \cap D \text{ modulo } Z_P(\mathbb{Z}_p)Z(\mathbb{Q}_p)\right\},$$

where Z_P (resp. Z) is the center of M_P (resp. G). Then $\{\xi_\alpha = \alpha(p)\}_{\alpha\in\Sigma}$ generates $D/T_G(\mathbb{Z}_p)Z(\mathbb{Q}_p)$, and $\mathcal{R} \cong \mathbb{Z}[T_G(\mathbb{Z}_p)Z(\mathbb{Q}_p)][T(\xi_\alpha)]_\alpha$, where as before $\mathbb{Z}[T_G(\mathbb{Z}_p)Z(\mathbb{Q}_p)]$ is the group algebra of the discrete group $T_G(\mathbb{Z}_p)Z(\mathbb{Q}_p)$. Similarly we have $\mathcal{R}_B \cong \mathbb{Z}[\frac{Z(\mathbb{Q}_p)}{Z(\mathbb{Z}_p)}][T(\xi_\alpha)]_\alpha$. If $G = GL(n)$,

$$\Sigma = \{\alpha_j | \alpha_j(p) = \mathrm{diag}[1_j, p1_{n-j}]\ (j = 1, 2, \ldots, n-1)\},$$

and the cocharacter α_j corresponds to the two-block parabolic subgroup $P = \{\left(\begin{smallmatrix} a & * \\ 0 & d \end{smallmatrix}\right)\}$ with $j \times j$-matrices a (and $(n-j) \times (n-j)$-matrices d).

For $\boldsymbol{\xi} = \prod_{\alpha \in \Sigma} \xi_\alpha$, we have $N(\mathbb{Q}_p) = \cup_{j=0}^{\infty} \boldsymbol{\xi}^j N \boldsymbol{\xi}^{-j}$. Thus writing $T(\xi) = [N\xi N]$ for $\xi \in D$ as an operator on $V^N = H^0(N, V)$, we see easily from (5.4) that $T(\boldsymbol{\xi}^j) = T(\boldsymbol{\xi})^j$ and for each finite-dimensional subspace $X \subset V(B)$, $T(\boldsymbol{\xi})|_X$ is nilpotent on X^N by (5.4).

For any $\mathcal{R}$-eigenvector $v \in V^N$ with $\overline{v}t = \lambda(t)\overline{v}$ ($t \in T_G(\mathbb{Q}_p)$, $\overline{v} = v \bmod V(B)$), we get

$$v|[NxN] = [N : x^{-1}Nx]\lambda(x)v = |\det(Ad_N(x))|_p \lambda(x)v, \tag{5.5}$$

where $|\ |_p$ is the standard absolute value of $\mathbb{Q}_p$ such that $|p|_p^{-1} = p$ and $Ad = Ad_N$ is the adjoint representation of T_G on the Lie algebra of N.

Let us prove a canonical splitting $V^N = V_B \oplus V(B)^N$ as $\mathcal{R}$-modules. We have by definition, $V^N = V^{N(\mathbb{Z}_p)} = \bigcup_r V^{S_1(r)}$. The subspace $V_r = V^{S_1(r)}$ is finite-dimensional and stable under $\mathcal{R}$. By the Jordan decomposition applied to $T(\boldsymbol{\xi})$, we can decompose uniquely that $V_r = V_r^\circ \oplus V^{nil}$ so that $T(\boldsymbol{\xi})$ is an automorphism on V_r° and is nilpotent on V^{nil}. We may replace $T(\boldsymbol{\xi})$ by $T(\boldsymbol{\xi}^a) = T(\boldsymbol{\xi})^a$ for any positive a in the definition of the above splitting. Since $T(\boldsymbol{\xi})$ is nilpotent over any finite-dimensional subspace of $V(B)$, V_r° injects into V_B; so, $\dim V_r^\circ$ is bounded by $\dim V_B \le |\mathfrak{W}|$. For any T_G-eigenvector $\overline{v} \in V_B$, lift it to $v \in V$. Then for sufficiently large j, $v\pi(\boldsymbol{\xi}^{-j})$ is in V^N. Since $\overline{v}\pi(\boldsymbol{\xi}^{-j})$ is a constant multiple of $\overline{v}$, we may replace $\overline{v}$ and v by $\overline{v}\pi(\boldsymbol{\xi}^{-j})$ and $v\pi(\boldsymbol{\xi}^{-j})$, respectively. Then for sufficiently large k, $w = vT(\boldsymbol{\xi}^k) \in V_r^\circ$, and $\overline{w}T(\boldsymbol{\xi})^{-k}$ is equal to $\overline{v}$ for the image $\overline{w}$ in V_B. This finishes the proof when the action of T_G on V_B is semi-simple. In general, take a sufficiently large r so that V_r surjects down to V_B. We apply the above argument to the semi-simplification of V_r under the action of the Hecke algebra. Thus $V^\circ = \bigcup_r V_r^\circ \cong V_B$, and this finishes the proof of $V^N = V_B \oplus V(B)^{N(\mathbb{Z}_p)}$ as $\mathcal{R}$-modules.

5.1.3 Rational Representations of G

We now fix a finite extension K over $\mathbb{Q}_p$. Let W be the p-adic integer ring of K. We write $S = S_0(r)$ ($r > 0$). Recall the Iwahori decomposition $S = N' T_G(\mathbb{Z}_p) N$. We consider the space $\mathcal{C}(R) = \{\phi : S/N(\mathbb{Z}_p) \to R\}$ made up of continuous functions for $R = W$ and K. Here the continuity is with respect to the p-adic topology of the source and the target. We would like to make $\mathcal{C}$ a left Δ_p^{-1}-module for the opposite semi-group Δ_p^{-1} of $\Delta_p = \Delta_S$. This Δ_p^{-1}-module $\mathcal{C}$ can be considered as the p-adic continuous induction of the right regular representation of $T_G(\mathbb{Z}_p)$ on the space of p-adically continuous functions on $T_G(\mathbb{Z}_p)$ with values in R (thus, this is the third induction process we encounter in this section). For that, we first define a left action of Δ_p on $Y_S = S/N(\mathbb{Z}_p)$. Since S acts on $Y_S = S/N(\mathbb{Z}_p)$ from the left, we only need to

define a left action of D. Pick $yN(\mathbb{Z}_p) \in Y_S$ and by the Iwahori decomposition, we may assume that $y \in N'T_G(\mathbb{Z}_p) \subset S$ and consider yN. Then for $d \in D$, $dyNd^{-1} = dyd^{-1}dNd^{-1} \subset dyd^{-1}N(\mathbb{Q}_p)$ and $dyd^{-1}N(\mathbb{Q}_p)$ is well defined in $G(\mathbb{Q}_p)/N(\mathbb{Q}_p)$. Since conjugation by $d \in D$ expands $N(\mathbb{Z}_p)$ and shrinks N', we have $dyd^{-1} \in S$, and the coset $dyd^{-1}N(\mathbb{Q}_p) \cap S = dyd^{-1}S$ is a well-defined single coset of N, which we define to be the image of the action of $d \in D$. We now let Δ_p^{-1} act on $\mathcal{C}$ by $d\phi(y) = \phi(d^{-1}y)$. In this way, $\mathcal{C}$ becomes a Δ_p^{-1}-module. Note that this action factors through $D/Z(\mathbb{Q}_p)$; so, if the group G has an involution $d \mapsto d^\iota$ with $dd^\iota \in Z(\mathbb{Q}_p)$, $d\phi(y) = \phi(d^\iota y) = \phi(d^{-1}y)$.

We consider the algebrogeometric induction module:

$$L_G(\kappa; K) = \left\{\phi \in H^0(G/N, \mathcal{O}_{G/N}) \middle| \phi(yt) = \kappa(t)\phi(y)\ \forall t \in T_G\right\}, \tag{5.6}$$

where $\mathcal{O}_{G/N}$ is the structure sheaf of the scheme G/N, and we regard its global section ϕ as a function on $G(\mathbb{Q}_p)/N(\mathbb{Q}_p)$ with values in K. We let G act on $L(\kappa; K) = L_G(\kappa; K)$ by $g\phi(y) = \phi(g^{-1}y)$. Then $L(\kappa; K) = \mathrm{ind}_B^G \kappa^{-1}$ (following the normalization of induction as in [RAG] I.3.3), which is the induction in the category of scheme theoretic representations (i.e., rational representations). We call κ *dominant* if $L(\kappa; K) \neq 0$. We write a matrix form of this representation as $\rho_\kappa = \rho_\kappa^G : G \to GL(L(\kappa; K))$.

All dominant weights in $X(T_G) \cong \mathbb{Z}^r$ form a cone

$$C(\omega_1, \ldots, \omega_r) = \{m_1\omega_1 + \cdots + m_r\omega_r | 0 \leq m_j \in \mathbb{Z}\}$$

if G is semi-simple (see [RAG] II.6.6). Here the generators ω_j are called fundamental weights. Take a weight ω of T_G. Since the derived group G_1 of G is semi-simple, its restriction to $T = T_G \cap G_1$ can be written as $\omega|_T = m_1\omega_1 + \cdots + m_r\omega_r$ for fundamental weights ω_j of G_1. The weight ω is called *regular* if $m_j > 0$ for all j, and we sometimes call a weight ω *sufficiently regular* if $m_j \gg 0$ for all j (without specifying how large m_j is).

We restrict functions in $L(\kappa; K)$ to $Y_S = S/N(\mathbb{Z}_p)$ and regard $L(\kappa; K) \subset \mathcal{C}(K)$. Multiply $L(\kappa; K)$ by a character $\varepsilon : T_G(\mathbb{Z}/p^r\mathbb{Z}) = S_0(r)/S_1(r) \to W^\times$ (regarding it as a function in $\mathcal{C}(W)$). Since $\mathbb{Q}_p^\times = \mathbb{Z}_p^\times \times p^{\mathbb{Z}}$, we can decompose $T_G(\mathbb{Q}_p) = T_G(\mathbb{Z}_p) \times \left(p^{\mathbb{Z}}\right)^r$ for the rank r of T_G, and we can extend ε to $T_G(\mathbb{Q}_p)$ requiring it to be trivial on $\left(p^{\mathbb{Z}}\right)^r$. In this way, we get the twisted Δ_p^{-1}-module $L(\kappa\varepsilon; K) = \varepsilon L(\kappa; K) \subset \mathcal{C}$. The action preserves $L(\kappa\varepsilon; W) = L(\kappa\varepsilon; K) \cap \mathcal{C}(W)$ but the original ρ_κ may not be. For $\xi \in D$, the action of $\xi^{-1} \in \Delta_p^{-1}$ is given by $\kappa(\xi)^{-1}\varepsilon(\xi)\rho_\kappa(\xi^{-1})$. For a W-module R, we define $L(\kappa\varepsilon; R) = L(\kappa\varepsilon; W) \otimes_W R$, which is a Δ_p^{-1}-module under the modified action.

We recall that the p-component of the semi-group $\Delta_0(p\mathfrak{N})$ introduced in Section 4.2.5 is a subsemi-group of the semi-group Δ_p. The definitions of the Hecke operator in (4.89) and (5.8) below are compatible, because we let $\delta \in \Delta_0(p\mathfrak{N})^\iota$ act on $L(\kappa\varepsilon; K)$ by $x \mapsto \delta x$ in the definition of the Hecke operator (4.89) and $\delta\phi(y) = \phi(\delta^\iota y) = \phi(\delta^{-1}y)$ for the main involution ι of $GL(2)$ as already remarked. This modification (in order to preserve p-integrality)

does not show up for coherent sheaves for $G = SL(2)$, because $L_C(\kappa; K)$ is one-dimensional due to the fact that $C = T$ for $G = GL(2)$. A similar remark applies to $L_G(\kappa; K)$ when $G = G_1 = \mathrm{Res}_{F/\mathbb{Q}}SL(2)$. However, in the Hilbert modular case of $G = \mathrm{Res}_{F/\mathbb{Q}}GL(2)$, the induction of $\kappa^* = (\kappa_1, \kappa_2) \in X(T_G)$ for the diagonal torus T_G of G is isomorphic to the tensor product $\det^{\kappa_1} \otimes L(\kappa_2 - \kappa_1 - I; K)$ for the restriction $\kappa_2 - \kappa_1 - I$ of κ to $T = T_G \cap G_1$ ($G_1 = \mathrm{Res}_{F/\mathbb{Q}}SL(2)$). The p-integral action created through this process is the action on $L_G(\kappa_2 - \kappa_1 - I; K)$ removing the "det" factor.

5.1.4 Nearly p-Ordinary Representations

We now study global cohomological automorphic representations. Let $G(\widehat{\mathbb{Z}}) \subset G(\mathbb{A}^{(\infty)})$ denote a maximal compact subgroup hyperspecial everywhere (by abusing notation; see [Tt1] for hyperspecial maximal compact subgroups). We assume that the p-component of $G(\widehat{\mathbb{Z}})$ is given by $G(\mathbb{Z}_p)$. For split groups over $\mathbb{Q}$ (like $GL(n)$ and $GSp(2n)$), they are just maximal compact subgroups. We take $G(\widehat{\mathbb{Z}}) = GL_n(\widehat{\mathbb{Z}})$ if $G = GL(n)$ and $G(\widehat{\mathbb{Z}}) = GSp_{2g}(\widehat{\mathbb{Z}})$ if $G = GSp(2g)$.

Hereafter we assume that κ is dominant in $X(T_G)$ and write ε for a character of $T_G(\mathbb{Z}/p^r\mathbb{Z})$. Let S be an open subgroup of $G(\widehat{\mathbb{Z}})$. We consider the associated modular variety $X(S) = G(\mathbb{Q})\backslash G(\mathbb{A})/(S \cdot C_{\mathbf{0}}) \cong \bigsqcup_{\Gamma} X_{\Gamma}$, where Γ runs over finitely many discrete subgroups $G(\mathbb{Q}) \cap t(S \cdot G(\mathbb{R})^+)t^{-1}$ in $G(\mathbb{R})^+$ for a coset decomposition $G(\mathbb{A}) = \bigsqcup_t G(\mathbb{Q})t(S \cdot G(\mathbb{R}))$. For the symmetric space $X = G(\mathbb{R})/C_{\mathbf{0}}$, we have written X_{Γ} for $\Gamma\backslash X$. For any W-module R, we then define a right action of $u \in SC_{\mathbf{0}}$ on $L(\kappa\varepsilon; R)$ by $\phi|u = \rho_{\kappa\varepsilon}(u_p^{-1})\phi$ (assuming that $S_p \subset S_0(r)$).

We define the covering space $\mathcal{X}(S)$ of $X(S)$ by

$$\mathcal{X}(S) = G(\mathbb{Q})\backslash(G(\mathbb{A}) \times L(\kappa\varepsilon; R))/SC_{\mathbf{0}} \cong \bigsqcup_{\Gamma} \mathcal{X}_{\Gamma}, \tag{5.7}$$

where $\gamma(x, \phi)u = (\gamma xu, \phi|u)$ for $\gamma \in G(\mathbb{Q})$ and $u \in SC_{\mathbf{0}}$, and we form the quotient space $\mathcal{X}_{\Gamma} = \Gamma\backslash(X \times L(\kappa\varepsilon; R))$ by the diagonal action. We use the same symbol $L(\kappa\varepsilon; R)$ for the sheaf of locally constant sections of $\mathcal{X}(S)$ over $X(S)$. Writing $[\xi]$ for the action of ξ^{-1} on $L(\kappa\varepsilon; R)$ for $\xi \in \Delta_p^{-1} \times G(\mathbb{A}^{(p)})$, we define as in (4.89) (and (4.37))

$$[S\xi S] = \mathrm{Tr}_{S/(S^{\xi} \cap S)} \circ [\xi] \circ \mathrm{Res}_{S/{}^{\xi}S \cap S}. \tag{5.8}$$

The choice of the schematic induction action as the action of ξ_p^{-1} is standard in the theory of automorphic representations (with coefficients in a characteristic 0 field). We write $T(\xi)$ for the Hecke operator $[S\xi S]$ under this choice of the action. When we study the p-adic Hecke operators, we choose the modified p-integral action. We write $T_p(\xi)$ for the operator $[S\xi S]$ under this p-adically integral action of $\xi \in \Delta_S$. Note that the factor $|\det(g)|_{\mathbb{A}}$ in (4.37) is missing in this definition of topological Hecke operators. Actually this factor (which coincides with $|\det(Ad_N(\xi))|_p$ when $\xi \in \Delta_0(p^{\infty}\mathfrak{N})$ and $G = GL(2)$) is absorbed by the Eichler–Shimura map as we show later in (5.40).

We consider the limit, shrinking S,

$$\mathcal{L}(R) = \mathcal{L}^q(\kappa\varepsilon; R) = \varinjlim_S H^q_!(X(S), L(\kappa\varepsilon; R)). \tag{5.9}$$

Here $H^q_!(X(S), L(\kappa\varepsilon; R))$ for $R = W, K$ is the image of the compactly supported cohomology $H^q_c(X(S), L(\kappa\varepsilon; R))$ in $H^q(X(S), L(\kappa\varepsilon; K))$. The group $H^q_!(X(S), L(\kappa\varepsilon; R))$ is called the *interior* cohomology of $X(S)$ with coefficients in $L(\kappa\varepsilon; R)$.

On the space $\mathcal{L}(K)$, the group $G(\mathbb{A}^\infty)$ acts from the right via a smooth representation, which is completely reducible. In particular, we have an action on $H^0(S, \mathcal{L}^q(\kappa\varepsilon; K)) = \mathcal{L}^q(\kappa\varepsilon; K)^S$ of the double coset algebra

$$R(S, G(\mathbb{A}^{p\infty}) \times \Delta_p) \cong R(S^{(p)}, G(\mathbb{A}^{p\infty})) \otimes \mathcal{R}_B,$$

where $S = S_p \times S^{(p)}$ and we have assumed that $S_p = S_0(r)$.

For $\xi \in D$, $T(\xi)$ acts on $\mathcal{L}^q(\kappa\varepsilon; W)^N$ ($N = N(\mathbb{Z}_p)$) through the Δ_p^{-1}-module structure of $L(\kappa\varepsilon; W)$. We write this operator as $T_p(\xi)$. On the other hand, $T(\xi)$ acts on $\mathcal{L}^q(\kappa\varepsilon; K)^N$ through the action of $G(\mathbb{A}^\infty)$. The corresponding operator is written by the same symbol $T(\xi)$. Then the two operators $T_p(\xi)$ and $T(\xi)$ are related on the image of $\mathcal{L}^q(\kappa\varepsilon; W)^N$ by

$$T_p(\xi) = \kappa(\xi)^{-1} T(\xi). \tag{5.10}$$

This follows from the fact that the normalized action of $\xi^{-1} \in \Delta_p^{-1}$ is given by $\kappa(\xi)^{-1}\varepsilon(\xi)\rho_\kappa(\xi)^{-1}$, while as an element of $G(\mathbb{A}^{(\infty)})$, ξ^{-1} acts via $\varepsilon(\xi)\rho_\kappa(\xi)^{-1}$.

Suppose that G admits Shimura variety Sh_S of PEL type defined over its reflex field E (see Chapter 7 for generality of Shimura varieties), whose complex points give $X(S)$. When $\kappa = 0$ (the identity character), the action of the Hecke operator is motivic and is (canonically) induced by the Hecke correspondence $\mathcal{T}(\xi) \subset Sh_S \times Sh_S$, and in this case, $T_p(\xi) = T(\xi)$ is the action of the correspondence on the Betti cohomology of $X(S)$. If $\kappa > 0$, we may relate cohomology groups of the sheaf $L(\kappa\varepsilon; K)$ as a part of the cohomology group with constant coefficients of a certain self-product Z of copies of the universal abelian scheme over $X(S)$. The Hecke operator then has an interpretation as an isogeny action on the universal abelian scheme, and it can be regarded as the action induced by the Hecke correspondence in $Z \times Z$. The action of $T_p(\xi)$ and $T(\xi)$ uses different action of Δ_p^{-1}. This action of Δ_p^{-1} determines the part of the cohomology group over Z identified with the cohomology group over $X(S)$ with locally constant (but nonconstant) coefficients. Thus the motivic realization of the two operators $T(\xi)$ and $T_p(\xi)$ could be actually different, and the operator $T_p(\xi)$ may not even have any motivic realization (as in the Hilbert modular case of nonparallel weight). For example, in Scholl's construction [Sc] of the Grothendieck motive associated with an elliptic Hecke eigenform f, if one changes the action of congruence subgroup Γ of $SL_2(\mathbb{Z})$ by a power of determinant character, the physical sheaf over $X_\Gamma(\mathbb{C})$ obtained is the same, but its rational structure (including the

Galois action) is different. In the elliptic modular case (treated in [Sc]), we can construct the motive associated with the standard p-adic Galois representation ρ_f of f and its Tate twists $\rho_f(m)$ as the étale realization of motives directly realized over a self-product of the universal elliptic curve.

For a Hilbert–Hecke eigenform f, we could twist ρ_f locally at each p-adic place by a power of the p-adic cyclotomic character, but this twist may not extend to a global twist because the exponent of the cyclotomic character depends on the p-adic place. In particular, if f is of nonparallel weight, the process of defining $T_p(\xi)$ corresponds to untwisting ρ_f to reach a $\mathfrak{p}$-ordinary Galois representation at each p-adic place $\mathfrak{p}$, which cannot be performed globally; so, the operator $T_p(\xi)$ may not have a motivic interpretation.

Unfortunately, for a Hilbert–Hecke eigenform f for $F \neq \mathbb{Q}$, the motive of f realized from the Hilbert modular variety does not produce a rank 2 motive whose étale realization gives rise to the two-dimensional Galois representation of f (see [BlR], [Ta], and also [MFG] 5.6.1 for a description of such Galois representations). Therefore to obtain such a rank 2 motive, we need to resort to some tricky construction either moving to Shimura curves treated in Section 4.3 (see [H81]) or extracting a rank 2 motive from a rank 4 motive (which can be constructed from some Shimura surfaces; see [Bl]). Despite such an effort, this problem is not completely solved yet.

We take $\xi \in D$ such that $N(\mathbb{Q}_p) = \bigcup_\jmath \xi^\jmath N(\mathbb{Z}_p)\xi^{-\jmath}$. The product $\boldsymbol{\xi} = \prod_{\alpha\in\Sigma}\xi_\alpha$ (or its power $\prod_{\alpha\in\Sigma}\xi_\alpha^{e_\alpha}$ with $e_\alpha > 0$) has this property. As seen at the beginning of Section 3.3, the p-adic (or profinite) limit $e = \lim_{n\to\infty} T_p(\boldsymbol{\xi})^{n!}$ exists as an operator on $H^q(X(S), L(\kappa\varepsilon; R))$ for $R = W$ and K (and is independent of the choice of $\boldsymbol{\xi}$), because $H^q(X(S), L(\kappa\varepsilon; R))$ is an R-module of finite type. Thus the limit e exists as an endomorphism of $\mathcal{L}^q(\kappa\varepsilon; R)^N$ for $R = W$ and K. It is easy to see, if $S_p \supset N$,

$$H^0(S, e(\mathcal{L}^q(\kappa\varepsilon; K)^N)) = e\,(H^q_!(X(S), L(\kappa\varepsilon; K)))\,. \tag{5.11}$$

We write $\mathcal{L}^q_{n.ord}(\kappa\varepsilon; R)$ for $e\left(\mathcal{L}^q(\kappa\varepsilon; R)^N\right)$. An irreducible representation π of $G(\mathbb{A}^{\infty)})$, which is a subquotient of $\mathcal{L}^q(\kappa\varepsilon; K)$, is called *nearly p-ordinary* of p-type κ if $eV(\pi)^N \neq 0$ for the representation space $V(\pi)$ of π.

5.1.5 Semi-Simplicity of Interior Cohomology Groups

Let $\pi \cong \otimes_\ell \pi_\ell$ be a cohomological automorphic representation of p-type κ. Suppose that its p-component π_p (acting on $V(\pi_p)$) is a subquotient of $\mathrm{Ind}_B^G \widetilde{\lambda}$ (this is automatic if π is nearly p-ordinary). Then we find a character $\lambda : T_G(\mathbb{Q}_p) \to K^\times$ with the above property such that $V_B[\widetilde{\lambda}] \neq 0$ and

$$|\det(Ad_N(x))^{-1}\widetilde{\lambda}(x)|_p = \left||\det(Ad_N(x))|_p\widetilde{\lambda}(x)\right|_p \leq |\kappa(x)|_p.$$

The equality holds if and only if π is nearly p-ordinary (in this case, automatically $V_B \neq 0$ and $\mathrm{Ind}_B^G \widetilde{\lambda} \twoheadrightarrow \pi_p$ because $V^N \cong V_B \oplus V(B)^N$ as $\mathcal{R}$-modules).

For the moment, suppose that $G(\mathbb{Q}_p) = GL_n(\mathbb{Q}_p)$ and factor the character λ as $\lambda(\mathrm{diag}[t_1, \ldots, t_n]) = \prod_{i=1}^n \lambda_i(t_i)$. Define the Hecke polynomial (at p) by $H_\pi(X) = \prod_{i=1}^n (1 - \lambda_i(p)X)$, and write Δ_N for the Newton polygon of $H_\pi(X)$ (e.g., [NAZ]). Define the Hodge polygon Δ_H of π to be the Newton polygon of $\prod_{i=1}^n (1 - (\kappa\varepsilon\rho)_i(p)X)$. Then the above inequality implies $\Delta_N \geq \Delta_H$ and the two extreme ends of the two polygons match.

We return to a general group G and assume that π is nearly p-ordinary. By definition, $\int_{N(\mathbb{Q}_p)} \phi(u)du = \delta_B(b) \int_{N(\mathbb{Q}_p)} \phi(b^{-1}xb)du$. This shows that

$$\delta_B = |\det \circ Ad_N|_p^{-1}. \tag{5.12}$$

By definition, $2\rho = \det \circ Ad_N$ is a sum of positive roots, and ρ is a sum of fundamental weights with respect to B. This shows

$$|\lambda|_p = |\kappa\rho|_p. \tag{5.13}$$

Since the dominant κ is nonnegative with respect to B, $\kappa\rho > 0$; that is, $\kappa\rho$ is in the interior of the Weyl chamber of B. This shows that if $w \neq 1$,

$$|\lambda^w(d)|_p < |\lambda(d)|_p \quad \text{for all } d \in D, \tag{5.14}$$

because $\mathfrak{W}$ acts simply transitively on Weyl chambers and each element in the interior of the chamber of λ has the maximum p-adic absolute value in its conjugates under $\mathfrak{W}$. In particular, we get

Theorem 5.3 *Let π be an irreducible nearly p-ordinary representation of p-type κ. Then there exists a character $\lambda : T_G(\mathbb{Q}_p) \to K^\times$ such that $\widetilde{\lambda} \hookrightarrow V(\pi_p)_B \hookrightarrow \oplus_{w\in\mathfrak{W}} \widetilde{\lambda^w}$ and $|\lambda|_p = |\rho\kappa|_p$, where ρ is the sum of fundamental weights with respect to B and $|\ |_p$ is the absolute value on K. Moreover, $eH^0(N(\mathbb{Z}_p), V(\pi_p))$ is one-dimensional, on which $T(\xi) = S\xi S$ for $\xi \in D$ acts by scalar $|\rho(\xi)|_p\lambda(\xi)$.*

Now suppose that $S = S_p \times G(\widehat{\mathbb{Z}}^{(p)})$ with $S_p = S_0(r)$ for $r > 0$. From the above theorem, we get the following semi-simplicity of the Hecke algebra (for cohomological nearly p-ordinary cusp forms of p-type κ) from the fact that the spherical irreducible representation of $G(\mathbb{Q}_\ell)$ has a unique vector fixed by any (given) hyperspecial maximal compact subgroup (e.g., [BlR1] 1.12), because of the Iwasawa decomposition of $G(\mathbb{Q}_\ell)$.

Corollary 5.4 *Let the notation and the assumption be as above. Then the Hecke module $eH_!^q(X(S), L(\kappa\varepsilon; K))$ is semi-simple.*

The projector $e = e_p$ is defined over $\mathcal{L}^q(\kappa\varepsilon; \overline{\mathbb{Q}})^N$. Thus the above semi-simplicity remains true on $e_P H_!^q(X(S), L(\kappa\varepsilon; \overline{\mathbb{Q}}))$ for $e_P = \prod_{\ell\in P} e_\ell$ with a finite set P of primes (where G is quasi-split over $\mathbb{Z}_P$) and an open compact subgroup S hyperspecial outside P and Iwahori at $\ell \in P$. For such nearly P-ordinary automorphic forms, semi-simplicity as in Corollary 5.4 holds.

5.2 Explicit Symmetric Domains

We describe unitary Lie groups $U^\psi(\mathbb{R})$ associated with Hermitian forms ψ and their symmetric domain X^ψ in an elementary way. Since each connected component of X^ψ is isomorphic to $\mathcal{D} = U^\psi(\mathbb{R})/C^\psi$ for a maximal compact subgroup C^ψ of $U^\psi(\mathbb{R})$, we describe an explicit form of $\mathcal{D}$ as a classical Hermitian domain, which is useful in describing κ^* from a given $\kappa \in X(T)$. We often follow the treatment by Shimura in [EPE]. In this section, R_n^m for a ring R denotes the set of all $m \times n$ matrices with entries in R.

5.2.1 Hermitian Forms over $\mathbb{C}$

Let V be a d-dimensional vector space over $\mathbb{C}$. We suppose having a Hermitian form $\varphi : V \times V \to \mathbb{C}$. Thus we have

(U1) $\varphi(\lambda v + \mu w, x) = \lambda\varphi(v,x) + \mu\varphi(w,x)$ for $v, w, x \in V$ and $\lambda, \mu \in \mathbb{C}$;
(U2) $\varphi(x,y) = \overline{\varphi(y,x)}$ for $x, y \in V$.

A Hermitian form φ is called positive (resp. negative) semi-definite if $\varphi(x,x) \geq 0$ (resp. $\varphi(x,x) \leq 0$) for all $x \in V$. We write $\varphi \geq 0$ (resp. $\varphi \leq 0$) if φ is positive (resp. negative) semi-definite. The form φ is called positive definite if it is positive semi-definite and $\varphi(x,x) = 0$ implies $x = 0$. In this case, we write $\varphi > 0$. If $\varphi - \psi > 0$, we write $\varphi > \psi$. This is a partial ordering on the set of Hermitian forms on V. Similarly we define negative definiteness.

Take a base $\mathbf{v} = \{v_1, \ldots, v_d\}$ of V, and identify V with the d-dimensional row vector space $\mathbb{C}_d$ by $\iota_{\mathbf{v}} : v = \sum_j x_j v_j \mapsto (x_1, \ldots, x_d) \in \mathbb{C}_d$. We have $\varphi(v,w) = \sum_{i,j} x_i \overline{y}_j \varphi(v_i, v_j)$ for $\iota_{\mathbf{v}}(w) = (y_1, \ldots, y_d)$. Set $\Phi = \Phi_{\mathbf{v}} := (\varphi(v_i, v_j)) \in \mathbb{C}_d^d$. We get $\varphi(v,w) = \iota_{\mathbf{v}}(v)\Phi\iota_{\mathbf{v}}(w)^*$, where $X^* = {}^t\overline{X}$ for a matrix X with coefficients in $\mathbb{C}$. By (U2), we see $\Phi^* = \Phi$. Conversely, for a given $\Phi \in \mathbb{C}_d^d$ with $\Phi^* = \Phi$, if we define $\varphi(x,y) = x\Phi y^*$ for $x, y \in \mathbb{C}_d$, obviously φ is a Hermitian form. Thus the knowledge of Hermitian forms is equivalent to the knowledge of Hermitian matrices. We call Φ *positive definite* if $\varphi(x,y) = x\Phi y^*$ is positive definite. We write $S^d(\mathbb{C})$ for the set of Hermitian matrices of degree d and define $S_+^d(\mathbb{C}) \subset S^d(\mathbb{C})$ to be the subset of positive definite Hermitian forms.

When we change a base $\{v_1, \ldots, v_d\}$ into $\{v_1', \ldots, v_d'\}$, then we find a matrix $A = (a_{ij}) \in M_d(\mathbb{C})$ such that $v_i' = \sum_j a_{ij} v_j$. Thus $A\mathbf{v}' = \mathbf{v}$, where we have written $\mathbf{v}$ in the form of a column vector $\mathbf{v} = {}^t(v_1, \ldots, v_d)$. Since $\iota_{\mathbf{v}}(v)\mathbf{v} = v$, we see $\iota_{\mathbf{v}'}(v)A\mathbf{v} = \iota_{\mathbf{v}'}(v)\mathbf{v}' = v = \iota_{\mathbf{v}}(v)\mathbf{v}$ and we get $\iota_{\mathbf{v}'}(v)A = \iota_{\mathbf{v}}(v)$. This shows that $\iota_{\mathbf{v}'}(v)A\Phi_{\mathbf{v}}(\iota_{\mathbf{v}'}(w)A)^* = \iota_{\mathbf{v}}(v)\Phi_{\mathbf{v}}\iota_{\mathbf{v}}(w)^* = \varphi(v,w) = \iota_{\mathbf{v}'}(v)\Phi_{\mathbf{v}'}\iota_{\mathbf{v}'}(w)^*$, and $A\Phi_{\mathbf{v}}A^* = \Phi_{\mathbf{v}'}$.

For $x \in \mathbb{C}$, we write $\mathrm{Tr}(x) = x + \overline{x}$. We now compute Φ choosing a nicer base. Let $N = Rad(V) = \{v \in V | \varphi(v, V) = 0\}$. Obviously $N \subset V$ is a subspace. This space is called the *null space* or *radical* of V. We have $\varphi(v+n, w+m) = \varphi(v,w) + \varphi(v,m) + \varphi(n,w) + \varphi(n,m) = \varphi(v,w)$ for $n, m \in N$. Thus φ is well-defined on the quotient $\overline{V} = V/N$, on which $Rad(\overline{V}) = \{0\}$.

First suppose that $Rad(V) = 0$, which is equivalent to $\det \Phi \neq 0$. Then we define the unitary group of φ by

$$U^{\varphi}(\mathbb{R}) = \{T \in \mathrm{End}_{\mathbb{C}}(V) | \varphi(xT, yT) = \varphi(x, y)\ \forall x, y \in V\}$$
$$\cong \{T \in M_d(\mathbb{C}) | T\Phi_{\mathbf{v}} T^* = \Phi_{\mathbf{v}}\} = \{T \in M_d(\mathbb{C}) | T^* \Phi_{\mathbf{v}}^{-1} T = \Phi_{\mathbf{v}}^{-1}\}. \quad (5.15)$$

By definition, $U^{\varphi}(\mathbb{R}) = U^{\lambda\varphi}(\mathbb{R})$ for any $0 \neq \lambda \in \mathbb{R}$.

For every $x \in V$, there exists $y \in V$ such that $\varphi(x, y) \neq 0$. Then we have $\varphi(x + y, x + y) = \varphi(x, x) + \mathrm{Tr}(\varphi(x, y)) + \varphi(y, y)$. If $\mathrm{Tr}(\varphi(x, y)) = 0$, then $\overline{\varphi}(x, y) = -\varphi(x, y)$. Replacing y by iy with $i = \sqrt{-1}$, we may assume that $\mathrm{Tr}(\varphi(x, y)) \neq 0$. Thus at least one of $\{\varphi(x + y, x + y), \varphi(x, x), \varphi(y, y)\}$ is non-zero. Therefore, we can find v_1 with $\varphi(v_1, v_1) \neq 0$ and $\varphi(v_1, v_1) \in \mathbb{R}$ by (U2). Let $W = (\mathbb{C}v_1)^{\perp} = \{v \in V | \varphi(v, v_1) = 0\}$. Then W is a subspace of V not containing v_1; that is, $\dim W < \dim V$. For any $v \in V$, $w = v - \frac{\varphi(v, v_1)}{\varphi(v_1, v_1)} v_1 \in W$; $V = W + \mathbb{C}v_1$. Since $\dim W < \dim V$, this shows $V = W \oplus \mathbb{C}v_1$. By induction, we find a base $v_2, \ldots, v_d$ of W such that $\varphi(v_i, v_j) = 0$ if $i \neq j$ and $0 \neq \varphi(v_i, v_i)$. Reordering v_i, we may assume that $\varphi(v_1, v_1), \ldots, \varphi(v_m, v_m)$ are positive and $\varphi(v_{m+1}, v_{m+1}), \ldots, \varphi(v_{m+n}, v_{m+n})$ $(d = m + n)$ are negative. The numbers (m, n) are called the *signature* of φ. Since m is the dimension of maximal subspace of V on which φ is positive definite, (m, n) depends only on φ. Changing φ by $-\varphi$ if necessary (this does not change U^{φ}), we may assume that $m \geq n$. Hereafter we always assume this and write $m = n + t$ with $t \geq 0$.

When $Rad(V) \neq 0$, we take a base $\overline{\mathbf{v}} = \{\overline{v}_1, \ldots, \overline{v}_j\}$ as above of $\overline{V} = V/N$ and lift them to vectors $v_1, \ldots, v_j$ of V. Set $V_0 = \mathbb{C}v_1 + \cdots + \mathbb{C}v_j$. Then we have $V = N \oplus V_0$. Supplementing $\mathbf{v} = \{v_1, \ldots, v_j\}$ with a base $v_{j+1}, \ldots, v_n$ of N, we have $\Phi_{\mathbf{v}} = \begin{pmatrix} \Phi_{\overline{\mathbf{v}}} & 0 \\ 0 & 0 \end{pmatrix} \in S^n(\mathbb{C})$, where $\Phi_{\overline{\mathbf{v}}} \in M_j(\mathbb{C})$ for $j = \dim \overline{V}$.

Theorem 5.5 *We have the following assertions which are equivalent.*

(1) *Let φ be a Hermitian form on a complex vector space V of dimension d. Then there exist subspaces V_+, V_-, and N of V such that*

$$V = V_+ \oplus V_- \oplus N,\ \varphi(V_+, V_-) = 0,\ \varphi(N, V) = 0,$$

and φ is positive (resp. negative) definite on V_+ (resp. V_-). Moreover, we can find bases of V_+ and V_- so that $V_+ = \mathbb{C}v_1 + \cdots + \mathbb{C}v_m$ and $V_- = \mathbb{C}v_{m+1} + \cdots + \mathbb{C}v_{m+n}$, with $\varphi(v_i, v_i) = \pm 1$ and $\varphi(v_i, v_j) = 0$ for $i \neq j$.

(2) *Let Φ be a Hermitian matrix of degree d. Then there exists a matrix $X \in GL_d(\mathbb{C})$ such that $X\Phi X^* = \begin{pmatrix} A & 0 & 0 \\ 0 & -B & 0 \\ 0 & 0 & 0 \end{pmatrix}$ and A and B are positive definite of degrees m and n, respectively. Moreover, we can take X so that $A = 1_m$ and $B = 1_n$.*

Hereafter we assume that φ is nondegenerate (i.e., $\det \Phi \neq 0$). There is another realization of a given Hermitian form if it is isotropic. A Hermitian form is called *anisotropic* if $\varphi(v, v) = 0$ implies $v = 0$. If φ is not anisotropic,

it is called *isotropic*. A subspace I of V is called *totally isotropic* if $I^\perp \cap I = \{x \in I | \varphi(x, I) = 0\} = I$. There exists a maximal totally isotropic subspace of V (because $\dim V < \infty$). If φ is isotropic, the maximal totally isotropic subspace has positive dimension.

Let J be a subspace of V with $J \subset J^\perp$; that is, J is totally isotropic. We claim that we have a totally isotropic subspace J' such that $J' \cong \mathrm{Hom}_{\mathbb{C}}(J, \mathbb{C})$ by φ. We prove this by induction on $\dim J$. Let $0 \neq v_1 \in J$. Then as we have seen, we always have $w \in V$ such that $\mathrm{Tr}(\varphi(w, v_1)) \neq 0$. We consider the equation for $x \in \mathbb{R}$,

$$\begin{aligned} 0 = \varphi(w + xv_1, w + xv_1) &= \varphi(w, w) + x\mathrm{Tr}(\varphi(w, v_1)) + \varphi(v_1, v_1)x^2 \\ &= \varphi(w, w) + x\mathrm{Tr}(\varphi(w, v_1)). \end{aligned} \tag{5.16}$$

Since $\mathrm{Tr}(\varphi(w, v_1)) \neq 0$, we can solve this equation and define $v_1' = w + xv_1$ for the solution x. If $\dim J = 1$, we set $J' = \mathbb{C}v_1'$, which proves the claim. Suppose now that $\dim J > 1$. We take a subspace J_1 of J so that $\dim J_1 = \dim J - 1$. Since J is totally isotropic, J_1 is totally isotropic. By the induction assumption, we find J_1' so that φ induces the duality between J_1' and J_1. Then any $v \in J_1' \cap J_1$ is orthogonal to every $w \in J_1$; so, $J_1' \cap J_1 = 0$. We put $X = J_1' + J_1$, on which φ is nondegenerate. Thus $V = X \oplus X^\perp$. Let J_0 be the projection of J to $X^\perp$, which is totally isotropic and of dimension 1. We find inside $X^\perp$ a subspace J_0' dual to J_0 under φ, and $J' = J_1' \oplus J_0'$ does the job.

Theorem 5.6 *We have the following assertions.*

(1) *If J is a totally isotropic subspace of V, we can find a totally isotropic subspace J' of V such that φ induces a $\mathbb{C}$-antilinear isomorphism $i : J' \cong \mathrm{Hom}_{\mathbb{C}}(J, \mathbb{C})$. Here "$\mathbb{C}$-antilinearity" means that $i(\lambda v) = \overline{\lambda} i(v)$ for $\lambda \in \mathbb{C}$.*
(2) *If J is a totally isotropic subspace of V, we can find a base $\mathbf{v} = \{v_1, \ldots, v_d\}$ of V such that*
   (a) *v_i for $i \leq j$ span J;*
   (b) *$\varphi(v_i, v_{n-j+k}) = \delta_{ik}$ for $0 \leq k \leq j$;*
   (c) *$\varphi(v_i, v_k) = 0$ if $j < k \leq d - j$ and $i \leq j$ or $i > d - j$.*
(3) *If J is a totally isotropic subspace of V, we can find a base $\mathbf{v} = \{v_1, \ldots, v_d\}$ of V such that the v_is for $i \leq j$ span J, and $\Phi_{\mathbf{v}} = \begin{pmatrix} 0 & 0 & 1_j \\ 0 & \Psi & 0 \\ 1_j & 0 & 0 \end{pmatrix}$.*
(4) *If Φ is isotropic, then we find an integer $j > 0$ and $X \in GL_d(\mathbb{C})$ such that $X\Phi X^* = \begin{pmatrix} 0 & 0 & 1_j \\ 0 & \Psi & 0 \\ 1_j & 0 & 0 \end{pmatrix}$ for an anisotropic Ψ.*

5.2.2 Symmetric Spaces of Unitary Groups

Let $I_{m,n} = \begin{pmatrix} 1_m & 0 \\ 0 & -1_n \end{pmatrix}$, and set for $d = m + n$

$$U_{m,n} = U_{m,n}(\mathbb{R}) = \{x \in GL_d(\mathbb{C}) | x^* I_{m,n} x = I_{m,n}\} \tag{5.17}$$

and $\mathfrak{X} = \left\{X \in \mathbb{C}_d^d \middle| X^* I_{m,n} X = \left(\begin{smallmatrix} A & 0 \\ 0 & -B \end{smallmatrix}\right)\ A > 0, B > 0\right\}$, which is nonempty by Theorem 5.5. Then we can let $(g,h) \in GL_m(\mathbb{C}) \times GL_n(\mathbb{C})$ act $\mathfrak{X}$ on the right and $\alpha \in U_{m,n}$ on the left by $X \mapsto \alpha X \left(\begin{smallmatrix} g & 0 \\ 0 & h \end{smallmatrix}\right)$. We now prove

Proposition 5.7 *Let* $\mathcal{D}(m,n) = \{z \in \mathbb{C}_n^m | z^* z < 1_n\}$ *and* $X(z) = \left(\begin{smallmatrix} 1_m & z \\ z^* & 1_n \end{smallmatrix}\right)$ *for* $z \in \mathbb{C}_n^m$. *Then* $\mathcal{D}(m,n) \times GL_m(\mathbb{C}) \times GL_n(\mathbb{C}) \cong \mathfrak{X}$ *by* $(z,g,h) \mapsto X(z)\left(\begin{smallmatrix} g & 0 \\ 0 & h \end{smallmatrix}\right)$.

Proof. Write $X \in \mathfrak{X}$ as $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ with $a \in \mathbb{C}_m^m$. Then

$$X^* I_{m,n} X = \left(\begin{smallmatrix} a^*a - c^*c & a^*b - c^*d \\ b^*a - d^*c & b^*b - d^*d \end{smallmatrix}\right) = \left(\begin{smallmatrix} A & 0 \\ 0 & -B \end{smallmatrix}\right)$$

for positive definite A and B. Thus $a^*a > c^*c \geq 0$, because $x^*c^*cx = (cx)^*(cx) \geq 0$. Thus if $au = 0$ for $u \in \mathbb{C}^m$, then $u^*(a^*a)u = 0$ and hence $u = 0$. This shows the linear map $\mathbb{C}^m \to \mathbb{C}^m$ given by $u \mapsto au$ is injective and hence invertible. That is, $\det a \neq 0$. Similarly, $b^*b < d^*d$ implies that $\det d \neq 0$. Then we define $z = bd^{-1}$. The fact $d^*(d^{-*}b^* - ca^{-1})a = b^*a - d^*c = 0$ tells us $z^* = (bd^{-1})^* = ca^{-1}$. Thus $X = \left(\begin{smallmatrix} 1_m & z \\ z^* & 1_n \end{smallmatrix}\right)\left(\begin{smallmatrix} a & 0 \\ 0 & d \end{smallmatrix}\right) = X(z)\left(\begin{smallmatrix} a & 0 \\ 0 & d \end{smallmatrix}\right)$. Since $0 > b^*b - d^*d = d^{-*}(1 - z^*z)d$, we have $z^*z < 1_n \Leftrightarrow 1_m > zz^*$ because the signature of $I_{m,n}$ is (m,n). This shows that $z \in \mathcal{D}(m,n)$, and the map is surjective. Suppose $\left(\begin{smallmatrix} 1_m & z \\ z^* & 1_n \end{smallmatrix}\right) = X(z) = X(z')\left(\begin{smallmatrix} c & 0 \\ 0 & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} c & z'd \\ z^*c & d \end{smallmatrix}\right)$. Then $c = d = 1$ and hence $z = z'$, which shows the injectivity. □

We now show that $X(z)$ is invertible. It is enough to show that $X(z)\mathbf{x} = 0$ implies $\mathbf{x} = 0$ for $\mathbf{x} \in \mathbb{C}^d$. Write $\mathbf{x} = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ with $x \in \mathbb{C}^m$. Then we have $x + zy = 0$ and $z^*x + y = 0$. Replace y in the first formula by $y = -z^*x$. We get $x - zz^*x = (1_m - zz^*)x = 0$. Since $1_m - zz^* < 0$, $\det(1_m - zz^*) \neq 0$ and hence $x = 0$. Then $y = -z^*x = 0$.

We now define $z \mapsto \alpha(z)$ for $\alpha \in U_{m,n}$ and $z \in \mathcal{D}(m,n)$ by

$$\alpha X(z) = X(\alpha(z)) \left(\begin{smallmatrix} \overline{h(\alpha,z)} & 0 \\ 0 & j(\alpha,z) \end{smallmatrix}\right)$$

with $h(\alpha,z) \in GL_m(\mathbb{C})$ and $j(\alpha,z) \in GL_n(\mathbb{C})$. Since we have

$$\begin{aligned} X((\alpha\beta)(z)) \left(\begin{smallmatrix} \overline{h(\alpha\beta,z)} & 0 \\ 0 & j(\alpha\beta,z) \end{smallmatrix}\right) &= \alpha\beta X(z) \\ &= \alpha X(\beta(z)) \left(\begin{smallmatrix} \overline{h(\beta,z)} & 0 \\ 0 & j(\beta,z) \end{smallmatrix}\right) \\ &= X(\alpha(\beta(z))) \left(\begin{smallmatrix} \overline{h(\alpha,\beta(z))h(\beta,z)} & 0 \\ 0 & j(\alpha,\beta(z))j(\beta,z) \end{smallmatrix}\right) \end{aligned}$$

for $\alpha, \beta \in U_{m,n}$, we conclude

$$h(\alpha\beta,z) = h(\alpha,\beta(z))h(\beta,z) \quad \text{and} \quad j(\alpha\beta,z) = j(\alpha,\beta(z))j(\beta,z), \tag{5.18}$$

$$(\alpha\beta)(z) = \alpha(\beta(z)), \tag{5.19}$$

because $X(z)$ is invertible. If $\alpha = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, we have $\alpha X(z) = \left(\begin{smallmatrix} a+bz^* & az+b \\ c+dz^* & cz+d \end{smallmatrix}\right)$ and

$$h(\alpha, z) = \overline{b} \cdot {}^t z + \overline{a},\ j(\alpha, z) = cz + d,\ \alpha(z) = (az + b)(cz + d)^{-1}. \tag{5.20}$$

Let $\mathbf{0}$ be the zero matrix in $\mathcal{D}(m, n)$. Then $\alpha(\mathbf{0}) = \mathbf{0}$ implies that

$$\alpha I_{m,n} = I_{m,n} \begin{pmatrix} \overline{h(\alpha,\mathbf{0})} & 0 \\ 0 & j(\alpha,\mathbf{0}) \end{pmatrix},$$

since $X(\mathbf{0}) = 1_{m+n}$. Thus $b = c = 0$. Then $\alpha I_{m,n}\alpha^* = I_{m,n}$ implies $a \in U_m(\mathbb{R})$ and $d \in U_n(\mathbb{R})$ for the classical definite unitary group $U_n(\mathbb{R})$ of degree n: $U_n(\mathbb{R}) = \left\{x \in GL_n(\mathbb{C}) \middle| xx^* = 1_n\right\}$. Thus we have

$$C_{\mathbf{0}} = \{\alpha \in U_{m,n} | \alpha(\mathbf{0}) = \mathbf{0}\} \cong U_m(\mathbb{R}) \times U_n(\mathbb{R}) \tag{5.21}$$

via $\alpha = \left(\begin{smallmatrix} a & 0 \\ 0 & d \end{smallmatrix}\right) \mapsto (a, d) \in U_m \times U_n$.

For each $X \in \mathfrak{X}$, writing $X^* I_{m,n} X = \left(\begin{smallmatrix} A & 0 \\ 0 & -B \end{smallmatrix}\right)$, we choose $u \in GL_m(\mathbb{C})$ and $v \in GL_n(\mathbb{C})$ so that $A = u^*u$ and $B = v^*v$ (by Theorem 5.5). Then $X \left(\begin{smallmatrix} u^{-1} & 0 \\ 0 & v^{-1} \end{smallmatrix}\right) \in U_{m,n}$, and hence $U_{m,n}(\mathbb{R})$ acts transitively on the quotient $\mathfrak{X}/(GL_m(\mathbb{C}) \times GL_n(\mathbb{C}))$. We have $\mathcal{D}(m, n) \cong U_{m,n}(\mathbb{R})/(U_m(\mathbb{R}) \times U_n(\mathbb{R}))$ by Proposition 5.7, and $U_{m,n}(\mathbb{R})$ acts transitively on $\mathcal{D}(m, n)$; that is, for any $z \in \mathcal{D}(m, n)$, there exists $\alpha \in U_{m,n}$ such that $\alpha(\mathbf{0}) = z$. The symmetric domain $\mathcal{D}(m, n)$ is bounded in $\mathbb{C}_n^m$ and is called a *bounded realization* of $U_{m,n}/C_{\mathbf{0}}$.

Corollary 5.8 *We have $\mathcal{D}(m, n) \cong U_{m,n}(\mathbb{R})/C_{\mathbf{0}}$ via $g \mapsto g(\mathbf{0})$, where $g(z) = (az+b)(cz+d)^{-1}$ for $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ with $a \in \mathbb{C}_m^m$. The maximal compact subgroup $C_{\mathbf{0}}$ is made up of $\left(\begin{smallmatrix} u & 0 \\ 0 & v \end{smallmatrix}\right)$ with $u \in U_m(\mathbb{R})$ and $v \in U_n(\mathbb{R})$.*

Since $X(w)^* I_{m,n} X(z) = \left(\begin{smallmatrix} 1-wz^* & z-w \\ w^*-z^* & w^*z-1 \end{smallmatrix}\right)$, replacing z and w by $z+\Delta z$ and z, we get

$$\begin{aligned} \operatorname{diag}[{}^t h(g,z), j(g,z)^*] \begin{pmatrix} 1-g(w)g(z)^* & \Delta g(z) \\ \Delta g(z)^* & g(w)^* g(z) - 1 \end{pmatrix} \operatorname{diag}[h(g,z), j(g,z)] = \\ \operatorname{diag}[{}^t h(g,z), j(g,z)^*] X(g(z)) I_{m,n} X(g(z+\Delta z)) \operatorname{diag}[h(g,z), j(g,z)] \\ = X(z)^* g^* I_{m,n} g X(z+\Delta z) = \begin{pmatrix} 1-wz^* & \Delta z \\ \Delta z^* & w^*z-1 \end{pmatrix}. \end{aligned}$$

From this, we conclude, for $dz = (dz_{ij})_{i,j}$,

$$dg(z) = {}^t h(g, z)^{-1} dz \cdot j(g, z)^{-1}. \tag{5.22}$$

We now exhibit an *unbounded realization*. We consider

$$\varphi = \begin{pmatrix} 0 & 0 & -\imath 1_n \\ 0 & \theta & 0 \\ \imath 1_n & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \Phi \\ \Psi & 0 \end{pmatrix}$$

with $\Phi \in \mathbb{C}_m^m$. Since φ has signature (m, n) $(m \geq n)$ if θ is positive definite, it is equivalent to $I_{m,n}$, and hence

$$U^{\varphi}(\mathbb{R}) = \left\{x \in GL_{m+n}(\mathbb{C}) \middle| x\varphi x^* = \varphi\right\}$$
$$= \left\{x \in GL_{m+n}(\mathbb{C}) \middle| x^*\varphi^{-1}x = \varphi^{-1}\right\} \cong U_{m,n}.$$

The isomorphism is given by $U^{\varphi} \ni x \mapsto XxX^{-1}$ for $X \in GL_{m+n}(\mathbb{C})$ with $X\varphi X^* = I_{m,n}$. Let

$$\mathfrak{Y} = \left\{Y \in \mathbb{C}_n^n \middle| Y^*\varphi^{-1}Y = \left(\begin{smallmatrix} A & 0 \\ 0 & -B \end{smallmatrix}\right),\ A > 0,\ B > 0\right\}.$$

Write $Y \in \mathfrak{Y}$ as

$$Y = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & j \end{pmatrix} = \begin{pmatrix} v_1 & c \\ v_2 & f \\ w & j \end{pmatrix}$$

so that $a, j \in \mathbb{C}_n^n$. Computing $Y^*\varphi^{-1}Y$, the $n \times n$ matrix at the bottom right corner is $ij^*c + f^*\theta^{-1}f - ic^*j < 0$. Thus $0 \leq f^*\theta^{-1}f < i(c^*j - j^*c)$. If $cu = 0$, then $iu^*(c^*j - j^*c)u = 0$ and hence $u = 0$. Thus $c \in GL_r(\mathbb{C})$ and similarly $j \in GL_r(\mathbb{C})$. We put $x' = cj^{-1}$ and $y' = fj^{-1}$. Then $z' = \begin{pmatrix} x' \\ y' \end{pmatrix}$ satisfies $i((x')^* - x') > (y')^*\theta^{-1}y'$. Put

$$\mathfrak{Z} = \mathfrak{Z}(n, \theta) = \left\{z = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) \middle| x \in \mathbb{C}_n^n,\ y \in \mathbb{C}_n^{m-n},\ \sqrt{-1}(x^* - x) > y^*\theta^{-1}y\right\}.$$

We have $z' \in \mathfrak{Z}$ and $(v^*, w^*)\varphi^{-1}\left(\begin{smallmatrix} v \\ w \end{smallmatrix}\right) > 0$ for $v = \left(\begin{smallmatrix} v_1 \\ v_2 \end{smallmatrix}\right) \in \mathbb{C}_m^m$. We want to show that v is invertible. It is enough to show that $vu = 0$; then $u = 0$. We have $u^*(v^*, w^*)\varphi^{-1}\left(\begin{smallmatrix} v \\ w \end{smallmatrix}\right)u = u^*(w^*\Psi v + v^*\Phi w)u = u^*w^*\Psi vu + (vu)^*\Phi wu = 0$, because $vu = 0$. Thus $u = 0$.

Define s by $\begin{pmatrix} x^* & y^* \\ 0 & -i\theta \end{pmatrix}s = v$. Then we have

$$Y = \begin{pmatrix} x^* & y^* & x \\ 0 & -i\theta & y \\ p & q & 1_n \end{pmatrix}\left(\begin{smallmatrix} s & 0 \\ 0 & j \end{smallmatrix}\right).$$

Let $Z = \begin{pmatrix} x^* & y^* & x \\ 0 & -i\theta & y \\ p & q & 1_r \end{pmatrix}$. Then $Z^*\varphi^{-1}Z = \left(\begin{smallmatrix} A & 0 \\ 0 & -B \end{smallmatrix}\right)$ with $A > 0,\ B > 0$. Looking at the lower left corner, we have $ix^* - ix^*p = 0$ and hence $p = 1_n$ because x is invertible (as we have shown c and j are invertible). Similarly looking at the middle block of the lowest row, we get $iy^* - iy^* - ix^*q = 0$ and hence $q = 0$. Thus $Y = B(z)\left(\begin{smallmatrix} s & 0 \\ 0 & j \end{smallmatrix}\right)$ for

$$B(z) = \begin{pmatrix} x^* & y^* & x \\ 0 & -\imath\theta & y \\ 1_n & 0 & 1_n \end{pmatrix} \tag{5.23}$$

for $z = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) \in \mathfrak{Z}$.

Proposition 5.9 *Let φ be as above. Then $\mathfrak{Y} \cong \mathfrak{Z} \times GL_m(\mathbb{C}) \times GL_n(\mathbb{C})$ by $B(z)\left(\begin{smallmatrix} a & 0 \\ 0 & b \end{smallmatrix}\right) \leftrightarrow (z, a, b)$.*

Proof. The argument before the proposition shows the surjectivity of the map $\mathfrak{Z} \times GL_m(\mathbb{C}) \times GL_n(\mathbb{C}) \to \mathfrak{Y}$. Suppose $B(z) = B(z')\left(\begin{smallmatrix} c & 0 \\ 0 & d \end{smallmatrix}\right)$. Writing $\Theta = \left(\begin{smallmatrix} 0 & -\imath\theta \\ 1_n & 0 \end{smallmatrix}\right)$, we have

$$\begin{pmatrix} z^* & z \\ \Theta & 1_n \end{pmatrix} = B(z) = B(z') \begin{pmatrix} c & 0 \\ 0 & d \end{pmatrix} = \begin{pmatrix} (z')^* c & z'd \\ \Theta c & d \end{pmatrix}.$$

From this we conclude that $d = 1$ and $\Theta c = \Theta$. Since Θ is invertible, $c = 1$. Thus the map is injective. □

By the proposition, we define $\alpha(z) \in \mathfrak{Z}$, $\vartheta(\alpha, z)$, $\mu(\alpha, z)$ for $z \in \mathfrak{Z}$ and $\alpha \in U^{\varphi}(\mathbb{R})$ by the following formula,

$$\alpha B(z) = B(\alpha(z)) \begin{pmatrix} \overline{\vartheta(\alpha,z)} & 0 \\ 0 & \mu(\alpha,z) \end{pmatrix}.$$

Again, for $\alpha, \beta \in U^{\varphi}(\mathbb{R})$, we have, similarly to (5.18),

$$\vartheta(\alpha\beta, z) = \vartheta(\alpha, \beta(z))\vartheta(\beta, z) \quad \text{and} \quad \mu(\alpha\beta, z) = \mu(\alpha, \beta(z))\mu(\beta, z), \tag{5.24}$$

$$(\alpha\beta)(z) = \alpha(\beta(z)). \tag{5.25}$$

Corollary 5.10 *Choosing $T \in GL_d(\mathbb{C})$ with $T\varphi T^* = I_{m,n}$ (by Theorem 5.5), the map $T : \mathfrak{Z} \to \mathcal{D}(m,n)$ given by $TB(z) = X(T(z)) \begin{pmatrix} \overline{h(T,z)} & 0 \\ 0 & j(T,z) \end{pmatrix}$ with $h(T,z) \in GL_m(\mathbb{C})$ and $j(T,z) \in GL_n(\mathbb{C})$ is a holomorphic isomorphism. Writing $\mathbf{i} = \begin{pmatrix} i1_n \\ 0 \end{pmatrix} \in \mathfrak{Z}$, we have $U^{\varphi}(\mathbb{R})/C_{\mathbf{i}} \cong \mathfrak{Z}$ by $g \mapsto g(\mathbf{i})$, where $C_{\mathbf{i}} = \{g \in U^{\varphi}(\mathbb{R}) | g(\mathbf{i}) = \mathbf{i}\}$. Furthermore, if we choose T so that $T(\mathbf{i}) = \mathbf{0}$ (possible by transitivity), $\alpha \mapsto T\alpha T^{-1}$ induces an isomorphism $C_{\mathbf{i}} \cong U_m(\mathbb{R}) \times U_n(\mathbb{R})$.*

Proof. By Propositions 5.7 and 5.9, $z \mapsto T(z)$ is a well-defined isomorphism. By definition, $\alpha \mapsto T\alpha T^{-1}$ is a surjective isomorphism of Lie groups from $U^{\varphi}(\mathbb{R})$ onto $U_{m,n}$. In particular, we have $T(\alpha(z)) = T\alpha T^{-1}T(z)$ by definition. If we can prove the transitivity of the action of $U^{\varphi}(\mathbb{R})$ on $\mathfrak{Z}$, we get another proof of the transitivity of the action of $U_{m,n}$ on $\mathcal{D}(m,n)$. To show the transitivity under $U^{\varphi}(\mathbb{R})$, we define $\xi : \mathfrak{Z} \to S_+^m(\mathbb{C})$ and $\eta : \mathfrak{Z} \to S_+^n(\mathbb{C})$ by

$$B(z)^* \varphi^{-1} B(z) = \begin{pmatrix} \overline{\xi(z)} & 0 \\ 0 & \eta(z) \end{pmatrix}. \tag{5.26}$$

By definition, we have

$$\mu(\alpha, w)^* \eta(\alpha(w)) \mu(\alpha, w) = \eta(w). \tag{5.27}$$

We want to find α with $\alpha(\mathbf{i}) = z$ for a given z. We assume that the last n rows of α are $(0, 0, d)$ with $d \in GL_n(\mathbb{C})$. Then $\alpha B(w) = B(\alpha(w)) \begin{pmatrix} * & 0 \\ 0 & d \end{pmatrix}$ because the lower right corner of $B(w)$ is 1_n. Thus $\mu(\alpha, w) = d$. Since $B(\mathbf{i})^* \varphi^{-1} B(\mathbf{i}) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$, we have $\eta(\mathbf{i}) = 2$ and $d^* \eta(z) d = 2$. Since $\eta(z) > 0$ by the definition of $\mathfrak{Z}$, we can find $a \in GL_r(\mathbb{C})$ such that $i(x^* - x) - y^* \theta^{-1} y = \eta(z) = 2aa^*$. Then for $z = {}^t({}^t x, {}^t y) \in \mathfrak{Z}$,

$$\alpha = p(z) = \begin{pmatrix} 1_n & iy^*\theta^{-1} & x - iaa^* \\ 0 & 1_{m-n} & y \\ 0 & 0 & 1_n \end{pmatrix} \begin{pmatrix} a & 0 & 0 \\ 0 & 1_{m-n} & 0 \\ 0 & 0 & a^{-*} \end{pmatrix} \in U^{\varphi}(\mathbb{R}) \tag{5.28}$$

satisfies $\alpha(\mathbf{i}) = z$, because

$$\alpha B(\mathbf{i}) = \begin{pmatrix} xa^{-*}-2\imath a & y^* & xa^{-*} \\ ya^{-*} & -\imath\theta & ya^{-*} \\ a^{-*} & 0 & a^{-*} \end{pmatrix} = B(z)\,\mathrm{diag}[\begin{pmatrix} a^* & 0 \\ -\imath\theta^{-1}y & 1_{m-n} \end{pmatrix}, a^{-*}].$$

This shows transitivity. The last isomorphism follows from Corollary 5.8. □

We define

$$\delta(z) = 2^{-n}\det(\eta(z)). \tag{5.29}$$

The factor 2^{-n} in the above definition is a normalization factor to make $\det(\delta(\mathbf{i})) = 1$. We then have by (5.27) the following transformation formula,

$$\delta(\alpha(z)) = \delta(z)|\det(\mu(\alpha, z))|^{-2} \quad (\text{for } \alpha \in U^{\varphi}(\mathbb{R})). \tag{5.30}$$

Let J be a totally isotropic subspace of V. We define

$$P(\mathbb{R}) = P_J^{\varphi}(\mathbb{R}) = \{g \in U^{\varphi}(\mathbb{R}) | gJ \subset J\}.$$

Writing $\varphi = \Phi_{\mathbf{v}}$ and taking $I = \sum_{j=m+1}^{m+n} \mathbb{C}v_j$, we see easily that $p(z) \subset P_I^{\varphi}(\mathbb{R})$. Thus the above proof of Corollary 5.10 shows

Corollary 5.11 *We have $U^{\varphi}(\mathbb{R}) = P_I^{\varphi}(\mathbb{R})C$ for any maximal totally isotropic subspace $I \subset V$ and any maximal compact subgroup C of $U^{\varphi}(\mathbb{R})$.*

The fact that $U^{\varphi}(\mathbb{R}) = P_J^{\varphi}(\mathbb{R})C$ holds for any totally isotropic subspace J.

Proof. The assertion follows from Corollary 5.10 for $C = C_{\mathbf{i}}$. Since any other C is a conjugate of $C_{\mathbf{i}}$, the assertion is clear. This fact follows from the topological fact that any compact subgroup has a fixed point on $U^{\varphi}(\mathbb{R})/C_{\mathbf{i}} = \mathfrak{Z}$. □

5.2.3 Invariant Measure

We now compute the measure on $\mathfrak{Z}$ invariant under the action of $U^{\varphi}(\mathbb{R})$. For that, we first compute several formulas:

$$B(z_1)^*\varphi^{-1}B(z) = \begin{pmatrix} \overline{\xi(z_1,z)} & \imath(z-z_1) \\ \imath(z-z_1)^* & -\eta(z_1,z) \end{pmatrix} \tag{5.31}$$

with $\xi(z_1, z) = \begin{pmatrix} \imath(\overline{x}_1 - {}^t x) & -\imath\cdot{}^t y \\ \imath\overline{y}_1 & \overline{\theta} \end{pmatrix}$ and $\eta(z_1, z) = i(x_1^* - x) - y_1^*\theta^{-1}y$. Thus $\eta(z) = \eta(z, z) = i(x^* - x) - y^*\theta^{-1}y$. We have by computation

$$B(z)\begin{pmatrix} 1 & 0\ 0 \\ \imath\theta^{-1} & 1\ 0 \\ -1 & 0\ 1 \end{pmatrix} = \begin{pmatrix} -\imath\eta(z) & y^* & x \\ 0 & -\imath\theta & y \\ 0 & 0 & 1 \end{pmatrix}.$$

From this we get

$$\det(B(z)) = \det(-i\theta)\det(-i\eta(z)). \tag{5.32}$$

Since we have $\alpha B(z) = B(\alpha(z))\begin{pmatrix} \overline{\vartheta(\alpha,z)} & 0 \\ 0 & \mu(\alpha,z) \end{pmatrix}$, we get

$$\det(\alpha)\det(B(z)) = \det(B(\alpha(z)))\overline{\det(\vartheta(\alpha,z))}\det(\mu(\alpha,z)). \tag{5.33}$$

From $\mu(\alpha,z)^*\eta(\alpha(z))\mu(\alpha,z) = \eta(z)$, we have for $\alpha \in U^{\varphi}(\mathbb{R})$,

$$\begin{aligned}
&\det(\alpha)\det(-i\theta)\det(-i\eta(z)) \\
&\quad = \det(\alpha)\det(B(z)) = \det(B(\alpha(z)))\det(\overline{\vartheta(\alpha,z)})\det(\mu(\alpha,z)) \\
&\quad = \det(-i\theta)\det(-i\eta(\alpha(z)))\det(\overline{\vartheta(\alpha,z)})\det(\mu(\alpha,z)) \\
&= \det(-i\theta)\det(-i\eta(z)\det(\mu(\alpha,z))^{-1}\det(\overline{\mu(\alpha,z)})^{-1}\det(\overline{\vartheta(\alpha,z)})\det(\mu(\alpha,z)) \\
&\quad = \det(-i\theta)\det(-i\eta(z))\det(\overline{\mu(\alpha,z)})^{-1}\det(\overline{\vartheta(\alpha,z)}).
\end{aligned}$$

This combined with $\det(\alpha)^{-1} = \overline{\det(\alpha)}$ shows

$$\det(\alpha)\det(\vartheta(\alpha,z)) = \det(\mu(\alpha,z)). \tag{5.34}$$

By replacing z and z_1 by $\alpha(z)$ and $\alpha(z_1)$, we get from

$$\alpha B(z) = B(\alpha(z))\begin{pmatrix} \overline{\vartheta(\alpha,z)} & 0 \\ 0 & \mu(\alpha,z) \end{pmatrix}$$

that

$$ {}^t\vartheta(\alpha,z_1)(\alpha(z)-\alpha(z_1))\mu(\alpha,z) = z - z_1, \tag{5.35}$$

$$\mu(\alpha,z_1)^*\eta(\alpha(z_1),\alpha(z))\mu(\alpha,z) = \eta(z_1,z), \tag{5.36}$$

$$\vartheta(\alpha,z_1)^*\xi(\alpha(z_1),\alpha(z))\vartheta(\alpha,z) = \xi(z_1,z). \tag{5.37}$$

To compute the Jacobian matrix of $z \mapsto \alpha(z)$, in (5.35), replace z_1 by $z+dz$. We get $\alpha^*dz = {}^t\vartheta(\alpha,z)^{-1}dz\mu(\alpha,z)^{-1}$. Writing

$$d\mathbf{z} = \bigwedge_{h=1}^{r+t}\bigwedge_{k=1}^{r} dz_{hk}\wedge d\overline{z}_{hk},$$

we have

$$\begin{aligned}
|\alpha^*d\mathbf{z}| &= |\det(\vartheta(\alpha,z)\otimes\mu(\alpha,z))|^{-2}|d\mathbf{z}| \\
&= |\det(\vartheta(\alpha,z))|^{-2n}|\det(\mu(\alpha,z))|^{-2m}|d\mathbf{z}| = |j(\alpha,z)|^{-2n-2m}|d\mathbf{z}|.
\end{aligned}$$

Since $\delta(\alpha(z)) = \delta(z)|j(\alpha,z)|^{-2}$, we see that $\delta(z)^{-m-n}|d\mathbf{z}|$ is the invariant measure.

Proposition 5.12 *The measure* $\delta(z)^{-m-n}(\otimes_{h=1}^{m}\otimes_{k=1}^{n}|dz_{hk}\otimes d\overline{z}_{hk}|)$ *gives an invariant measure on* $\mathfrak{Z}$ *under the action of* $U^{\varphi}(\mathbb{R})$.

5.3 The Eichler–Shimura Map

Before starting a detailed study of the nearly ordinary part of coherent cohomology groups, we make explicit a generalized Eichler–Shimura map for unitary and symplectic groups and hence the association of the weight: $\kappa \mapsto \kappa^*$ so that $H^0_{cusp}(X_\Gamma,\underline{\omega}^{\kappa}) \hookrightarrow H^d(X_\Gamma, L(\kappa^*;\mathbb{C}))$.

5.3.1 Unitary Groups

Recall $U = U_{m,n}(\mathbb{R}) = \left\{g \in GL_{m+n}(\mathbb{C}) \middle| gI_{m,n}g^* = I_{m,n}\right\}$, where $I_{m,n} = \left(\begin{smallmatrix} 1_m & 0 \\ 0 & -1_n \end{smallmatrix}\right)$. We have made explicit the quotient space $U/C_{\mathbf{0}}$ as the bounded symmetric domain $\mathcal{D} = \mathcal{D}(m,n)$. As we have seen, the complexification C of $C_{\mathbf{0}}$ is $GL_m(\mathbb{C}) \times GL_n(\mathbb{C})$. The functions ${}^t h^{-1}$ and j correspond to the standard representation of $GL(m)$ and $GL(n)$, respectively.

For the upper-triangular subgroup $B_m \subset GL(m)$, we write T_m for the diagonal torus in B_m. For a dominant weight κ of the diagonal torus $T = T_{m,n} = T_m \times T_n$ of $U_{m,n}$, writing ρ_κ for the rational representation into $GL(L_C(\kappa;\mathbb{C}))$ of $GL(m) \times GL(n)$ of weight κ, we have the automorphic factor

$$U_{m,n}(\mathbb{R}) \times \mathcal{D}(m,n) \ni (g,z) \mapsto \rho_\kappa({}^t h^{-1}(g,z), j(g,z)),$$

with which we associate a vector bundle $\underline{\omega}^\kappa$ on $\mathcal{D}(m,n)$ as we described in Section 1.1. Take a reductive group G admitting the Shimura variety whose symmetric domain X^+ is analytically isomorphic to a product $\prod_{\sigma\in\Sigma} \mathcal{D}(m_\sigma, n_\sigma)$ (for an index set Σ). Let T be the torus of G giving rise to the product $\prod_{\sigma\in\Sigma} T_{m_\sigma,n_\sigma}$. We have the vector bundle $\bigotimes_\sigma \underline{\omega}^{\kappa_\sigma}$ over X^+ associated with $\kappa \in X(T)$. For a congruence subgroup $\Gamma \subset G(\mathbb{Q})^+$, we want to embed $H^0(X_\Gamma, \underline{\omega}^\kappa)$ into $H^q(X_\Gamma, L_G(\kappa^*;\mathbb{C}))$ as Hecke modules for a suitable rational representation $L_G(\kappa^*;\mathbb{C})$ of G with highest weight $\kappa^* \in X(T)$. By the Künneth formula (cf. [CDR] I.1.3 and VII.5.2), $\kappa^* = \sum_\sigma \kappa_\sigma^*$, so we may assume that $|\Sigma| = 1$; thus, we drop the subscript σ from our notation.

We have seen in (5.34) that $\det(h(g,z)) = \det(g)^{-1}\det(j(g,z))$. This also can be shown as follows. On the diagonal torus $T_C \subset U(m) \times U(n)$, for $g = \mathrm{diag}[t_1,\dots,t_m,t_{m+1},\dots,t_{m+n}]$, t_j satisfies $\bar{t}_j = t_j^{-1}$ and $j(g,z) = cz+d = \mathrm{diag}[t_{m+1},\dots,t_{m+n}]$ (resp. $h(g,z) = \overline{a} + \overline{b}\cdot {}^t z = \mathrm{diag}[t_1^{-1},\dots,t_m^{-1}]$). The representation $g \mapsto {}^t h(g,\mathbf{0})^{-1}$ (resp. $g \mapsto j(g,\mathbf{0})$) corresponds to the standard representation of $GL_n(\mathbb{C})$ (resp. the standard representation of $GL_m(\mathbb{C})$); so, the corresponding highest character, after applying "det", is

$$\mathrm{diag}[t_1,\dots,t_{m+n}] \mapsto \prod_{j=1}^{m} t_j \quad (\text{resp. } \prod_{j=m+1}^{m+n} t_j).$$

This relation well explains the above formula (5.34). We thus embed the product $U(m) \times U(n)$ into $GL(m) \times GL(n)$ by $g \mapsto J(g) = ({}^t h(g,\mathbf{0})^{-1}, j(g,\mathbf{0}))$. We also write $J(g,z) = ({}^t h(g,z)^{-1}, j(g,z))$.

Writing $\mathbf{d}z = \bigwedge_{i,j} dz_{ij}$, we get from (5.22) that

$$\mathbf{d}g(z) = \det(g)^n \det(j(g,z))^{-m-n}\mathbf{d}z.$$

Write $\mu_{m,n} \in X(T)$ for the character

$$\mu_{m,n}(\mathrm{diag}[t_1,t_2,\dots,t_{m+n}]) = \left(\prod_{j=1}^{m} t_j\right)^{-n} \times \left(\prod_{k=1}^{n} t_{m+k}\right)^{m}. \tag{5.38}$$

Suppose that $\kappa \geq \mu_{m,n}$, and write $\kappa^* = \kappa - \mu_{m,n}$. We consider a polynomial function $p_{\kappa^*} : \mathcal{D} \to \mathrm{Hom}_{\mathbb{C}}(L_C(\kappa^*;\mathbb{C}), L_G(\kappa^*;\mathbb{C}))$ in z such that

$$p(\alpha(z)) \circ \rho^C_{\kappa^*}(J(\alpha,z)) = \alpha p(z) \quad (\rho^C_{\kappa^*} = \mathrm{Ind}^C_{B\cap C}\, \kappa^*)$$

for all $\alpha \in G$, where $C = GL(m) \times GL(n)$. Since $\mathcal{D} = G/C_{\mathbf{0}}$, if it exists, such a function is unique. We can define it by $p(\alpha(\mathbf{0})) \circ \rho^C_{\kappa^*}(J(\alpha,\mathbf{0})) = \alpha p(\mathbf{0})$. If we change α by αu for $u \in U(m) \times U(n)$, then we have

$$\begin{aligned} p(\alpha(\mathbf{0})) \circ \rho^C_{\kappa^*}(J(\alpha)J(u)) &= \alpha u \cdot p(\mathbf{0}) \\ \iff p(\alpha(\mathbf{0})) \circ \rho^C_{\kappa^*}(J(\alpha)) &= \alpha u \cdot p(\mathbf{0}) \circ \rho^C_{\kappa^*}(J(u))^{-1}. \end{aligned}$$

Such a map $p(\mathbf{0})$ with $up(\mathbf{0}) \circ \rho^C_{\kappa^*}(J(u))^{-1} = p(\mathbf{0})$ exists because the group $GL(m) \times GL(n)$ is identified with a subgroup of $GL_d(\mathbb{C}) = U(d)(\mathbb{C})$ $(d = m+n)$; thus, it corresponds to the inclusion

$$L_C(\kappa^*;\mathbb{C}) = \rho^C_{\kappa^*} \hookrightarrow \mathrm{ind}^{GL(m+n)}_P \rho^C_{\kappa^*} = L_G(\kappa^*;\mathbb{C}) = \rho^G_{\kappa^*}$$

for $P = \mathrm{diag}[GL(m), GL(n)]B$. Take $\kappa^*|_{T_C}$ to be the highest weight ω_n associated with the standard representation of $GL(n)$. Then κ^* corresponds to the standard representation of $U_{m,n}$, and we have $p_{\omega_n}(z)(x) = \left(\begin{smallmatrix} z \\ 1 \end{smallmatrix}\right) x$ for $x \in \mathbb{C}^n$. We easily verify that $gp_{\omega_n}(z)(x) = p_{\omega_n}(z)(j(g,z)x)$. Thus $p(z)$ is a polynomial in z in this special case. Similarly to the above, if $\kappa^*|_{T_C} = \omega_m$ corresponds to the contragredient of the standard representation of $GL(m)$, then κ^* is associated with the complex conjugate of the standard representation of $U_{m,n}$, and we have $p_{\omega_m}(z)(x) = \left(\begin{smallmatrix} 1_m \\ {}^t z \end{smallmatrix}\right) x$ for $x \in \mathbb{C}^m$. Again we verify that $\overline{g}p_{\omega_m}(z)(x) = p_{\omega_m}(z)(h(g,z)x)$, and $p_{\omega_m}(z)$ is a polynomial in z. For general κ, $L_C(\kappa^*;\mathbb{C})$ (resp. $L_G(\kappa^*;\mathbb{C})$) is a quotient of $L_C(\omega_n;\mathbb{C})^{\otimes t} \bigotimes L_C(\omega_m;\mathbb{C})^{\otimes s}$ (resp. $L_G(\omega_n;\mathbb{C})^{\otimes t} \bigotimes L_G(\omega_m;\mathbb{C})^{\otimes s}$). The general $p_{\kappa-\mu}(z)$ is a constant multiple of the projected image of the tensor product of copies of $p_{\omega_1}(z)$ and hence is a polynomial in z.

For a congruence subgroup $\Gamma \subset G(\mathbb{R})$, each global section of $\underline{\omega}^\kappa$ is a holomorphic function $f : \mathcal{D} \to L_C(\kappa;\mathbb{C})$ satisfying $f(\gamma(z)) = J_\kappa(\gamma,z)f(z)$ for $J_\kappa(\gamma,z) = \rho_\kappa(J(\gamma,z))$. We define for $f \in H^0(X_\Gamma, \underline{\omega}^\kappa_{/\mathbb{C}})$ a holomorphic differential with values in $L_G(\kappa^*;\mathbb{C})$ by $\omega(f) = p_{\kappa^*}(z)(f)\mathbf{d}z$. Note that here $L_C(\kappa;\mathbb{C}) = L_C(\kappa^*;\mathbb{C}) \otimes L_C(\mu_{m,n};\mathbb{C})$ and $L_C(\mu_{m,n};\mathbb{C})$ is one-dimensional; so, we can identify $L_C(\kappa;\mathbb{C})$ with $L_C(\kappa^*;\mathbb{C})$ canonically as vector spaces and, thus, the above definition is consistent. We verify that $\alpha^*\omega(f) = \rho^G_{\kappa^*}(\alpha)\omega(f)$.

Theorem 5.13 *Let Γ be a discrete subgroup of $U(m,n)(\mathbb{R})$ with the compact quotient space $X_\Gamma = \Gamma\backslash\mathcal{D}(m,n)$. Suppose that $\kappa \geq \mu_{m,n}$ as above, and define $\kappa^* = \kappa - \mu_{m,n}$. Then the association $f \mapsto [\omega(f)] \in H^q(X_\Gamma, L(\kappa^*;\mathbb{C}))$ for $q = \dim_{\mathbb{C}} \mathcal{D}$ induces an embedding $H^0(X_\Gamma, \underline{\omega}^\kappa) \hookrightarrow H^q(X_\Gamma, L_G(\kappa^*;\mathbb{C}))$, where $[\omega(f)]$ is the de Rham cohomology class of $\omega(f)$.*

When Γ is a noncocompact congruence subgroup of a global unitary group $G_{/\mathbb{Q}}$, the same assertion is valid replacing $\underline{\omega}^\kappa$ by its cuspidal subsheaf $\underline{\omega}_\kappa$.

The theorem follows from standard harmonic analysis when X_Γ is compact (see [CDR], [MaM] and also [LFE] 6.2 for the simplest case of $G = SL(2)_{/\mathbb{Q}}$). Even in the noncompact case, the same argument holds for cusp forms, since cuspidal sections decrease rapidly towards the boundary of the canonical (topological) compactification of X_Γ. The compactification is called the Borel–Serre compactification which is a manifold with corners (see [BoS]). Since the detailed proof of this fact relies on harmonic analysis, which is beyond the scope of this algebraic book, we admit this result.

5.3.2 Symplectic Groups

We give a brief description of the Siegel modular version of Theorem 5.13. Let $\mathfrak{H}_g = \{z = {}^t z \in \mathbb{C}_g^g | \operatorname{Im}(z) > 0\}$ be the Siegel upper half space, and take $G = GSp(2g)_{/\mathbb{Q}}$ for simplicity, although this works well for any reductive group $G_{/\mathbb{Q}}$ admitting Shimura varieties whose symmetric domain is isomorphic to a product of copies of $\mathfrak{H}_g$. Thus

$$G(R) = \left\{\alpha \in GL_{2g}(R) \middle| \alpha J_g \cdot {}^t\alpha = \nu(\alpha) J_g \ \text{ with } \nu(\alpha) \in R^\times\right\}.$$

For $z, w \in \mathfrak{H}_g$, we write $Y(z,w) = \left(\begin{smallmatrix} z & w \\ 1 & 1 \end{smallmatrix}\right)$, and for $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in GSp(2g)$, we have $gY(z,w) = Y(g(z), g(w)) \left(\begin{smallmatrix} cz+d & 0 \\ 0 & cw+d \end{smallmatrix}\right)$. From this, we get

$$\begin{aligned} \left(\begin{smallmatrix} 0 & w-z \\ z-w & 0 \end{smallmatrix}\right) &= \nu(g)^{-1} {}^t Y(z,w) \cdot {}^t g J g Y(z,w) \\ &= \nu(g)^{-1} \left(\begin{smallmatrix} {}^t(cz+d) & 0 \\ 0 & {}^t(cw+d) \end{smallmatrix}\right) \left(\begin{smallmatrix} 0 & g(w)-g(z) \\ g(z)-g(w) & 0 \end{smallmatrix}\right) \left(\begin{smallmatrix} cz+d & 0 \\ 0 & cw+d \end{smallmatrix}\right), \end{aligned}$$

which implies $w - z = \nu(g)^{-1} {}^t(cz+d)(g(w) - g(z))(cw+d)$. Replacing w by $z + dz$ for $dz = (dz_{ij})_{i,j}$, we get $d(g(z)) = \nu(g) {}^t(cz+d)^{-1} dz (cz+d)^{-1}$. We put $\mathbf{d}z = \bigwedge_{i \geq j} dz_{ij}$ and define a character $\mu : GL(g) \times \mathbb{G}_m \to \mathbb{G}_m$ by

$$\mathbf{d}g(z) = \mu(cz + d, \nu(g))^{-1} \mathbf{d}z. \tag{5.39}$$

Regard μ as a character of the maximal split torus $T_{GSp(2g)}$ of $GSp(2g)$ by the isomorphism $T_{GL(g)} \times \mathbb{G}_m \cong T_{GSp(2g)}$ given by $(t, t') \mapsto \operatorname{diag}[t, t' t^{-1}]$.

For a given character $\kappa > 0$ of the maximal split torus $T_{GL(g)}$, the automorphic factor $Sp_{2g}(\mathbb{R}) \times \mathfrak{H}_g \ni (g, z) \mapsto \rho_\kappa(J(g,z))$ with $J(\gamma, z) = cz + d$ for $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ gives rise to a vector bundle $\underline{\omega}^\kappa$ on $\mathfrak{H}_g$. Now suppose that κ is a character of $T_{GSp(2g)}$, and use the same symbol to denote its restriction to $T_{Sp(2g)} = T_{GL(g)}$. Suppose also that $\kappa \geq \mu$. Let $\kappa^* = \kappa - \mu$ and consider the $GSp(2g)$-module $L(\kappa^*; \mathbb{C})$ on which $t \in T_{GSp(2g)}$ acts by $\phi(g) \mapsto \phi(gt)$. We want to have a polynomial function $p : \mathfrak{H}_g \to \operatorname{Hom}_{\mathbb{C}}(L_{\mathbb{C}}(\kappa^*; \mathbb{C}), L(\kappa^*; \mathbb{C}))$ in coordinates of $z \in \mathfrak{H}_g$ such that $p(g(z)) \circ \rho_{\kappa-\mu}(J(g,z)) = gp(z)$ $(g \in Sp_{2g}(\mathbb{R}))$. We show the existence and uniqueness (up to a scalar multiple) of the map. If we change g by gu for $u \in C$ for the stabilizer C of $\mathbf{i} = i1_g$ $(i = \sqrt{-1})$ of $\mathfrak{H}_g$ in $Sp_{2g}(\mathbb{R})$, then the expected formula tells us

$$p(g(\mathbf{i}))\circ\rho_{\kappa-\mu}(J(g,\mathbf{i})J(u,\mathbf{i}))=gup(\mathbf{i})$$
$$\iff p(g(\mathbf{i}))\circ\rho_{\kappa-\mu}(J(g,\mathbf{i}))=gup(\mathbf{i})\circ\rho_{\kappa-\mu}(J(u,\mathbf{i}))^{-1}.$$

The homomorphism $u\mapsto J(u,\mathbf{i})$ identifies C with the definite unitary group $U(g)\subset GL_g(\mathbb{C})$. Such a map $p(\mathbf{i})$ with $up(\mathbf{i})\circ\rho_{\kappa-\mu}(J(u,\mathbf{i}))^{-1}=p(\mathbf{i})$ exists because $GL(g)$ is identified with a subgroup of $Sp(2g)(\mathbb{C})$; thus, it corresponds to the identity inclusion $\rho_{\kappa-\mu}\hookrightarrow \mathrm{Ind}_P^{GSp(2g)}\rho_{\kappa^*}=L(\kappa^*;\mathbb{C})$ for the Siegel parabolic subgroup $P\subset GSp(2g)$ with reductive part $GL(g)\times\mathbb{G}_m$. Then we extend $p(\mathbf{i})$ to $p(z)$ by $p(z)=gp(\mathbf{i})\circ\rho_{\kappa-\mu}(J(g,\mathbf{i}))^{-1}$, taking $g\in Sp_{2g}(\mathbb{R})$ with $z=g(\mathbf{i})$. From the construction of p, it is unique up to a constant multiple. Take $\kappa^*=\kappa-\mu$ to be the highest weight ω_1 associated with the standard representation of $GL(g)$. Then κ^* corresponds to the standard representation of $Sp(2g)$, and we have $p(z)(x)=\left(\begin{smallmatrix}z\\1\end{smallmatrix}\right)x$ for $x\in\mathbb{C}^g$. We easily verify that $gp(z)(x)=p(z)(J(g,z)x)$. Thus $p(z)$ is a polynomial in z in this special case. For general κ, $L_C(\kappa^*;\mathbb{C})$ and $L_G(\kappa^*;\mathbb{C})$ are quotients of the tensor product of several copies of $L_C(\omega_1;\mathbb{C})$ and $L_G(\omega_1;\mathbb{C})$, respectively. The function $p(z)$ for general κ is given by the projected image of the tensor product of copies of $p(z)$ for ω_1 and hence is again a polynomial in z.

Again for a congruence subgroup Γ of $GSp_{2g}(\mathbb{R})$, we may identify a global section of $\underline{\omega}_{\kappa/\mathbb{C}}$ over X_Γ with a holomorphic function $f:\mathfrak{H}_g\to L_C(\kappa;\mathbb{C})$ with $f(\gamma(z))=\rho_\kappa^C(J(\gamma,z))f(z)$ for all $\gamma\in\Gamma$. Then we define a differential with values in $L_G(\kappa^*;\mathbb{C})$ by $\omega(f)=p(z)(f)\mathbf{d}z$. Note that $L(\mu;\mathbb{C})$ is one-dimensional; so, we can identify $L(\kappa;\mathbb{C})$ with $L(\kappa^*;\mathbb{C})$ canonically (just multiplying by the character μ) as vector spaces, and thus the above definition is consistent. We verify from the above property characterizing $p(z)$ that $\gamma^*\omega(f)=\gamma\omega(f)$ for all $\gamma\in\Gamma(N)$. Thus for cuspidal holomorphic forms, the harmonic analysis in [CDR] again gives us

Theorem 5.14 *Let Γ be a congruence subgroup of $Sp_{2g}(\mathbb{Q})$ and μ be the weight of $T_{GSp(2g)}$ defined as above. Suppose that $\kappa\geq\mu$, and define $\kappa^*=\kappa-\mu$. Then the association sending f to $[\omega(f)]$ in the de Rham cohomology group $H^q(X_\Gamma,L(\kappa^*;\mathbb{C}))$ for $q=\dim_{\mathbb{C}}\mathfrak{H}_g$ induces the embedding $H^0(X_\Gamma,\underline{\omega}_\kappa)\hookrightarrow H^q(X_\Gamma,L_G(\kappa^*;\mathbb{C}))$, where $[\omega(f)]$ is the de Rham cohomology class of $\omega(f)$, and $\underline{\omega}_\kappa\subset\underline{\omega}^\kappa$ is the cuspidal subsheaf of $\underline{\omega}^\kappa$.*

We admit this fact. A more thorough comparison theorem (with proof) of de Rham cohomology and the Betti cohomology (with coefficients in locally constant sheaves) of Siegel modular varieties can be found in [DAV] Chapter VI.

5.3.3 Hecke Equivariance

We show that the Eichler–Shimura map is equivariant under Hecke operators and is compatible with our normalization of Hecke operators. We assume $G_{/\mathbb{Q}}$ to be a reductive group with $G_1(\mathbb{R})$ isomorphic either to $\prod_{\sigma\in\Sigma}SU_{m_\sigma,n_\sigma}(\mathbb{R})$ or to $\prod_{\sigma\in\Sigma}Sp_{2n_\sigma}(\mathbb{R})$ for a finite index set Σ, where $SU_{m,n}(\mathbb{R})$ is the derived

group of $U_{m,n}(\mathbb{R})$. In the former case, we call G unitary (or of type A) and in the latter case, we call G symplectic (or of type C). Again by the Künneth formula, to describe the equivariance, we may assume $|\Sigma| = 1$; so, we drop the subscript σ. We write μ for $\mu_{m,n}$ given in (5.38) when G is unitary and for μ in (5.39) when G is symplectic.

We have normalized the Hecke operator on the topological Betti cohomology group taking the action of $\xi \in \Delta_B$ normalized as $\widetilde{\rho}_\kappa^G(\xi^{-1}) = \kappa^{-1}(\xi)\rho_\kappa^G(\xi^{-1})$. Note that $\widetilde{\rho}_\kappa = \widetilde{\rho_\kappa \otimes \chi} = \widetilde{\rho}_{\chi\kappa}$ for an algebraic character $\chi : G$ or $C \to \mathbb{G}_m$.

We normalize again in the same way the action on $\underline{\omega}^\kappa$ taking the action of $\widetilde{\rho}_\kappa^C(\xi^{-1}) = \kappa^{-1}(\xi)\rho_\kappa^C(\xi^{-1})$ in addition to the division by $\mu(\xi)$. Let $\mathbb{T} = T_p(\xi)$ and also write coset representatives as ξ; that is, $T(\xi) = \bigsqcup_\xi N(\mathbb{Z}_p)\xi$. Recalling $\kappa^* = \kappa\mu^{-1}$, we have

$$\begin{aligned}\omega(f)|\mathbb{T}_B &= \sum_\xi (\kappa^*(\xi))^{-1}\rho_{\kappa^*}^G(\xi^{-1})p(\xi(z))(f(\xi(z)))\mathbf{d}(\xi(z)) \\ &= \sum_\xi p(z)((\kappa\mu^{-1}(\xi))^{-1}\rho_{\kappa-\mu}(J(\xi,z))^{-1})f(\xi(z))\mu(\xi)^{-1}\mathbf{d}z \\ &= p(z)\left(\mu(\xi)^{-1}\sum_\xi \kappa^{-1}(\xi)\rho_\kappa(J(\xi,z))^{-1}f(\xi(z))\right)\mathbf{d}z = \omega(f|\mathbb{T}_C). \quad (5.40)\end{aligned}$$

Here we have added the subscripts B and C in order to emphasize the space on which the operator acts; that is, B implies the topological Betti cohomology $H^q(X(U), L_G(\kappa^*;\mathbb{C}))$ and C indicates the coherent cohomology $H^0(Sh_U, \underline{\omega}^\kappa)$ (which is a part of the de Rham cohomology). In short, the extra modification of the action of the Hecke operator $T(\xi)$ by the character μ on the coherent cohomology is absorbed by $\mathbf{d}(\xi(z)) = \mu(\xi)^{-1}\mathbf{d}z$ in the topological cohomology. Hence the normalization of Hecke operators at p is identical on the left-hand side and the right-hand side of the Eichler–Shimura map. This is why we do not have such a factor in (5.8). We also remark that in the case of $GL(2)$, the determinant factor $|\det(g)|_{\mathbb{A}}$ in the definition (4.37) of the Hecke operator on the coherent cohomology is the value $\mu(g)^{-1}$ as above. Thus we do not have this factor in (4.89).

6

Moduli Schemes

We recall the construction of moduli of abelian schemes. The theory of moduli varieties of abelian varieties has been studied mainly by Shimura and Mumford in the years 1950s to 1960s. Shimura proved the existence of the moduli varieties over a canonically determined number field relative to a given endomorphism ring, a level N-structure and a polarization in the late 1950s to the early 1960s. This gives a modulus over the integer ring of the field with a sufficiently large number of primes inverted.

Basically at the same time, Grothendieck studied the moduli of subschemes in a given projective scheme $X_{/S}$ (flat over S) and also that of the Picard functors. The existence of a moduli scheme, the Hilbert scheme $\mathrm{Hilb}_{X/S}$, of closed flat subschemes of $X_{/S}$ enabled Mumford, via his theory of geometric quotients of quasi-projective schemes ([GIT]), to construct the moduli of abelian schemes with level N-structure over $\mathbb{Z}[\frac{1}{N}]$.

We briefly recall here the construction of Grothendieck and Mumford, limiting to the case which we need later. After that, we recall the construction of Shimura varieties with a canonical family of abelian varieties in the following chapter.

6.1 Hilbert Schemes

We describe the theory of the Hilbert scheme that classifies all closed S-flat subschemes of a given projective variety $X_{/S}$. This is a generalization of the earlier theory of Chow coordinates that classifies cycles on a projective variety. The theory is due to Grothendieck and the main source of our exposition is his Exposé 221 in [FGA]. One can find an exposition in [NMD] 8.2 on this subject under milder assumptions.

6.1.1 Vector Bundles

We recall a construction in [EGA] I.9.4 of algebrogeometric vector bundles over a scheme S. Let S be a Noetherian scheme and $\mathcal{E}$ be a locally free sheaf on S of rank n. For an open subscheme $O \subset S$, we consider the tensor algebra $\mathbf{T}(\mathcal{E}(O))$, which is a noncommutative graded $\mathcal{O}_S(O)$-algebra whose degree m component is $\mathbf{T}_m(\mathcal{E}(O)) = \overbrace{\mathcal{E}(O) \otimes_{\mathcal{O}_S(O)} \mathcal{E}(O) \otimes \cdots \otimes_{\mathcal{O}_S(O)} \mathcal{E}(O)}^{m}$ and $\mathbf{T}_0(\mathcal{E}(O)) = \mathcal{O}_S(O)$. The multiplication is induced by tensor product $(x, y) \mapsto x \otimes y$ for $x, y \in \mathcal{E}(O)$. Then $\mathbf{S}(\mathcal{E}(O))$ is the quotient of $\mathbf{T}(\mathcal{E}(O))$ by the homogeneous two-sided ideal generated by $x \otimes y - y \otimes x$ for all $x, y \in \mathcal{E}(O)$. The (commutative) $\mathcal{O}_S(O)$-graded algebra $\mathbf{S}(\mathcal{E}(O))$ is called the symmetric algebra of $\mathcal{E}(O)$. The functor $O \mapsto \mathbf{S}(\mathcal{E}(O))$ is a contravariant functor from the category of open subsets of S into graded algebras; so, it is a presheaf of algebras over S. We write $\mathbf{S}(\mathcal{E})$ for the sheafication of the presheaf of algebras ([GME] 1.1.2). Then by definition, $\mathbf{S}(\mathcal{E})$ is a graded $\mathcal{O}_S$-algebra. We define the relative spectrum $\mathbf{V}(\mathcal{E}) = \mathit{Spec}_S(\mathbf{S}(\mathcal{E}))$ as an S-scheme (e.g., [GME] 1.5.4).

When $S = \mathit{Spec}(\mathcal{B})$ is affine, for every multiplicative subset $\mathcal{S}$ of $\mathcal{B}$, it is plain that $\mathbf{S}(\mathcal{S}^{-1}M) = \mathcal{S}^{-1}\mathbf{S}(M)$ for $M = \Gamma(S, \mathcal{E})$. This shows that $\mathbf{S}(\mathcal{E})$ is the quasi-coherent sheaf $\widetilde{\mathbf{S}(M)}$ associated with the $\mathcal{O}_S$-module $\mathbf{S}(M)$ (cf. [GME] 1.2.3). For any $\mathcal{B}$-algebra R, a $\mathcal{B}$-linear map $\ell : M \to R$ extends uniquely to a $\mathcal{B}$-algebra homomorphism $\ell : \mathbf{S}(M) \to R$ by sending $x_1 \otimes x_2 \otimes \cdots \otimes x_m$ to $\ell(x_1)\ell(x_2)\cdots\ell(x_m)$. Thus the algebra $\mathbf{S}(M)$ represents the covariant functor $ALG_{/\mathcal{B}} \to AB$ given by $R \mapsto \mathrm{Hom}_{\mathcal{B}}(M, R)$, where $ALG_{/\mathcal{B}}$ (resp. AB) is the category of $\mathcal{B}$-algebras (resp. abelian groups). Since $ALG_{/\mathcal{B}}$ is antiequivalent to the category of affine S-schemes $AFF_{/S}$, $\mathbf{V}(\widetilde{M})$ represents the contravariant functor $AFF_{/S} \to AB$ given by $O \mapsto \mathrm{Hom}_{\mathcal{O}_S}(\widetilde{M}, \mathcal{O}_O) = \mathrm{Hom}_{\mathcal{O}_O}(\widetilde{M}_O, \mathcal{O}_O)$, where $M_O = M \otimes_{\mathcal{B}} R$ for $O = \mathit{Spec}(R)$; in other words, $\widetilde{M}_O = f^*\widetilde{M}$ for the structure morphism $f : O \to S$. See Section 4.4.1 in the text and [GME] 1.4 and 1.5 for representability of functors.

In general, covering S by affine open subsets $\{O\}_O$, we see that the restriction of $\mathbf{S}(\mathcal{E})$ to each affine open subset is quasi-coherent and, hence, $\mathbf{S}(\mathcal{E})$ is quasi-coherent over S. By further shrinking the affine open subsets O, we may assume that $\mathcal{E}|_O \cong \mathcal{O}_O^n$. Then $\mathbf{S}(\mathcal{E}(O))$ is isomorphic to the polynomial ring over $\mathcal{O}_S(O) = R$ generated by a base $T_1, \ldots, T_n$ of $\mathcal{E}(O)$ over R. This shows that $\mathbf{V}(\mathcal{E})$ is an S-scheme of finite type and hence Noetherian. Since $\mathbf{V}(\mathcal{E}|_O)$ represents the functor $S' \mapsto \mathrm{Hom}_{\mathcal{O}_O}(\mathcal{E}|_O, \mathcal{O}_{S'}) = \mathrm{Hom}_{\mathcal{O}_{S'}}(f^*(\mathcal{E}|_O), \mathcal{O}_{S'})$ for $f : S' \to S$, the schemes $\{\mathbf{V}(\mathcal{E}|_O)\}_O$ glue over S to represent the functor $S' \mapsto \mathrm{Hom}_{\mathcal{O}_S}(\mathcal{E}, \mathcal{O}_{S'}) = \mathrm{Hom}_{\mathcal{O}_{S'}}(\mathcal{E}_{S'}, \mathcal{O}_{S'})$ for $\mathcal{E}_{S'} = f^*\mathcal{E}$ (see Remark 1.5.1 on page 31 of [GME]); in other words, we have, for $f : S' \to S \in SCH_{/S}$,

$$\mathrm{Hom}_S(S', \mathbf{V}(\mathcal{E})) = \mathrm{Hom}_{\mathcal{O}_{S'}}(\mathcal{E}_{S'}, \mathcal{O}_{S'}) = \Gamma(S', \widehat{\mathcal{E}}_{S'}), \tag{6.1}$$

where $\widehat{\mathcal{E}}$ is the $\mathcal{O}_S$-dual of $\mathcal{E}$. The last equality in (6.1) follows from the definition of $\widehat{\mathcal{E}}$: $\mathrm{Hom}_{\mathcal{O}_{S'}}(\mathcal{E}_{S'}, \mathcal{O}_{S'}) = \Gamma(S', \widehat{\mathcal{E}}_{S'})$. Thus $\mathbf{V}(\mathcal{E})$ gives the vector bundle associated with the dual sheaf $\widehat{\mathcal{E}}$.

We consider a contravariant functor $S' \mapsto \mathrm{Hom}_{\mathcal{O}_{S'}}(\mathcal{E}_{S'}, \mathcal{F}_{S'})$ for $S' \xrightarrow{f} S$, where $\mathcal{E}$ and $\mathcal{F}$ are two locally free sheaves of finite rank. Note that

$$\mathrm{Hom}_{\mathcal{O}_{S'}}(\mathcal{E}_{S'}, \mathcal{F}_{S'}) \cong \widehat{\mathcal{E}}_{S'} \otimes_{\mathcal{O}_{S'}} \mathcal{F}_{S'} \cong f^*(\widehat{\mathcal{E}} \otimes_{\mathcal{O}_S} \mathcal{F}).$$

Then (6.1) tells us that the above functor is represented by $\mathbf{V}(\mathcal{E} \otimes_{\mathcal{O}_S} \widehat{\mathcal{F}})$:

$$\begin{aligned} \mathrm{Hom}_S(S', \mathbf{V}(\mathcal{E} \otimes_{\mathcal{O}_S} \widehat{\mathcal{F}})) &= \mathrm{Hom}_{\mathcal{O}_{S'}}(\mathcal{E}_{S'}, \mathcal{F}_{S'}) \\ &= \Gamma(S', f^*(\widehat{\mathcal{E}} \otimes \mathcal{F})) \quad \text{for } f : S' \to S \in SCH_{/S}. \end{aligned} \tag{6.2}$$

In particular, $S' \mapsto \mathrm{End}_{\mathcal{O}_{S'}}(\mathcal{E}_{S'})$ is represented by $\mathbf{V}(\mathcal{E} \otimes \widehat{\mathcal{E}})$.

Taking an affine open subset $O = Spec(R)$ so that $\mathcal{E} \cong \mathcal{O}_O^n$ and fixing an isomorphism $\mathcal{E} \cong \mathcal{O}_O^n$, each $\phi \in \mathrm{End}_{\mathcal{O}_{S'}}(\mathcal{E}_{S'})$ has matrix expression $\phi = (\phi_{ij})$. Then we may define $\det \phi \in \mathcal{O}_{S'}$. The construction of $\det \phi$ does not depend on the choice of the base of $\mathcal{E}|_O$; so, covering S by affine open subschemes, they glue together into a functorial homomorphism $\det : \mathrm{End}_{\mathcal{O}_{S'}}(\mathcal{E}_{S'}) \to \mathcal{O}_{S'}$ inducing a morphism $\det : \mathbf{V}(\mathcal{E} \otimes \widehat{\mathcal{E}})_{/S} \to \mathbf{G}_{a/S}$ by Yoneda's lemma (cf. [GME] Lemma 1.4.1). We define

$$GL(\mathcal{E})_{/S} = \det^*(\mathbf{G}_m) = \mathbf{V}(\mathcal{E} \otimes \widehat{\mathcal{E}}) \times_{\mathbf{G}_a} \mathbf{G}_m. \tag{6.3}$$

Then $GL(\mathcal{E})_{/S}$ is a group scheme (cf. [GME] 1.6) representing the functor $S' \mapsto \mathrm{Aut}_{\mathcal{O}_{S'}}(\mathcal{E}_{S'})$. In particular, we write $GL(n)_{/S} = GL(\mathcal{O}_S^n)$ and we have

$$GL(n)_{/\mathbb{Z}} = Spec(\mathbb{Z}[T_{ij}, \det(T_{ij})^{-1}]) \quad \text{and} \quad GL(n)_{/S} = GL(n) \times_{\mathbb{Z}} S. \tag{6.4}$$

We can slightly generalize the above argument. Let $\mathcal{F}$ be another locally free sheaf over S with the same rank as $\mathcal{E}$. We consider the functor from $SCH_{/S}$ into $SETS$ given by

$$\underline{\mathrm{Isom}}_S(\mathcal{E}, \mathcal{F})(S') = \left\{ \phi \in \mathrm{Hom}_{\mathcal{O}_{S'}}(\mathcal{E}_{S'}, \mathcal{F}_{S'}) \middle| \phi \text{ is an isomorphism} \right\}. \tag{6.5}$$

On a small open subscheme $O \subset S$ over which $\mathcal{E} \cong \mathcal{F} \cong \mathcal{O}_O^n$, the functor $\underline{\mathrm{Isom}}_S(\mathcal{E}|_O, \mathcal{F}|_O)$ is represented by an affine scheme $\mathrm{Isom}_S(\mathcal{E}|_O, \mathcal{F}|_O)$ isomorphic to $GL(\mathcal{E}|_O)$, identifying $\mathcal{F}|_O$ with $\mathcal{E}|_O$. Covering S by such open subschemes O, the schemes $\mathrm{Isom}_S(\mathcal{E}|_O, \mathcal{F}|_O)$ glue together into a scheme $\mathrm{Isom}_S(\mathcal{E}, \mathcal{F})$ affine over S by the uniqueness of the universal scheme. The resulting scheme $\mathrm{Isom}_S(\mathcal{E}, \mathcal{F})$ obviously represents the functor in (6.5).

6.1.2 Grassmannians

We define a contravariant functor $\mathrm{Grass} : SCH \to SETS$ by

$$\underline{\mathrm{Grass}}_{\mathcal{O}^n,m}(S) = \left[\pi : \mathcal{O}_S^n \twoheadrightarrow \mathcal{F} \middle| \pi \text{ surjective, } \mathcal{F} \text{ locally } \mathcal{O}_S\text{-free of rank } m\right].$$

As before, the straight brackets $[*]$ indicate the set of isomorphism classes of the objects $*$ inside the brackets. The quotient $\pi : \mathcal{O}_S^n \to \mathcal{F}$ is isomorphic to $\pi' : \mathcal{O}_S^n \to \mathcal{F}'$ if we have the following commutative diagram,

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathrm{Ker}(\pi) & \longrightarrow & \mathcal{O}_S^n & \stackrel{\pi}{\to} & \mathcal{F} & \longrightarrow & 0 \\ & & \| & & \| & & \wr\downarrow & & \\ 0 & \longrightarrow & \mathrm{Ker}(\pi') & \longrightarrow & \mathcal{O}_S^n & \stackrel{\pi'}{\to} & \mathcal{F}' & \longrightarrow & 0 \end{array}$$

with exact rows. For each morphism $f : S' \to S$, the pullback $f^*\pi : \mathcal{O}_{S'}^n \twoheadrightarrow f^*\mathcal{F}$ gives the contravariant functoriality. The stabilizer of $\pi : \mathcal{O}_S^n \to \mathcal{O}_S^m$ can be identified with the maximal parabolic subgroup

$$P = \left\{ \left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right) \in GL(n) \middle| d \text{ is of size } m \times m \right\}.$$

As is well known (see below Theorem 6.1), the quotient $\mathrm{Grass}_{\mathcal{O}^n,m} = GL(n)/P$ is a projective scheme defined over $\mathbb{Z}$. Intuitively, this projective scheme represents the functor $\underline{\mathrm{Grass}}_{\mathcal{O}^n,m}$, that is,

$$\mathrm{Hom}_{SCH}(S, \mathrm{Grass}_{\mathcal{O}^n,m}) \cong [\pi : \mathcal{O}_S^n \twoheadrightarrow \mathcal{F}]$$

functorially. We soon verify this fact in down-to-earth terms after stating the representability as a theorem (Theorem 6.1). Of course, if $m = n-1$ or 1, we have $\mathrm{Grass}_{\mathcal{O}^n,m} = \mathbf{P}^{n-1}$.

We can generalize this construction slightly. Let S be a scheme, and take a locally free sheaf $\mathcal{E}_{/S}$ of rank n. Then, for each S-scheme $S' \stackrel{f}{\to} S$, we define a contravariant functor from $S\text{-}SCH$ to $SETS$ by

$$\underline{\mathrm{Grass}}_{\mathcal{E},m}(S') = \left[\pi : \mathcal{E}_{S'} \twoheadrightarrow \mathcal{F} \middle| \pi \text{ surjective, } \mathcal{F} \text{ locally } \mathcal{O}_{S'}\text{-free of rank } m\right].$$

Covering S by sufficiently small open subschemes O_i so that $\mathcal{E}_{O_i} \cong \mathcal{O}_{O_i}^n$, we have $\underline{\mathrm{Grass}}_{\mathcal{O}_{O_i}^n,m}$ represented by $\mathrm{Grass}_{m/O_i} = \mathrm{Grass}_m \times O_i$. The gluing data $g_{ij} : \mathcal{O}_{O_i \cap O_j} \cong \mathcal{E}_{O_i \cap O_j} \cong \mathcal{O}_{O_j \cap O_i}$ give rise to a Čech cocycle g_{ij} with values in $GL(n)$. This gluing datum induces a gluing datum of $\left\{\mathrm{Grass}_{m/O_i}\right\}$, giving rise to the scheme $\mathrm{Grass}_{\mathcal{E},m}$ over S which should represent the above functor.

Theorem 6.1 *For each locally free sheaf $\mathcal{E}_{/S}$ of finite rank n, the functor $\underline{\mathrm{Grass}}_{\mathcal{E},m}$ is represented by a projective scheme $\mathrm{Grass}_{\mathcal{E},m}$ over S.*

Here we use the word "projective" in a strict sense: "strongly projective" in Section 6.1.6 that a scheme X over S is projective if we have a closed immersion over S of X into a projective space $\mathbf{P}^N \times_{\mathbb{Z}} S$ for a positive integer N, where $\mathbf{P}^N = Proj(\mathbb{Z}[X_0, \ldots, X_N])$. Thus if $S = Spec(\mathcal{B})$ and $\mathbb{G} = \mathrm{Grass}_{\mathcal{E},1} \cong \mathbf{P}^{n-1}$ for $n = \mathrm{rank}\,\mathcal{E}$, $\mathcal{O}_{\mathbb{G}}(1) = \mathcal{E}$ has to be free of rank n over $\mathcal{B}$; so, $\mathcal{E}$ has to be free. However, we can always find a surjective morphism of sheaves $\mathcal{O}_S^{N+1} \twoheadrightarrow \mathcal{E}$ if S is Noetherian, which induces a closed immersion $\mathrm{Grass}_{\mathcal{E},1} \hookrightarrow$

$\mathrm{Grass}_{\mathcal{O}_S^{N+1},1} = \mathbf{P}^N$; so, by enlarging the dimension of the projective space, we confirm that $\mathrm{Grass}_{\mathcal{E},1}$ is projective. We reduce in the following proof of projectivity of $\mathrm{Grass}_{\mathcal{E},m}$ to this case of $m = 1$ by taking Plücker's coordinates (described in the proof).

The scheme $\mathrm{Grass}_{\mathcal{E},m}$ gives a geometric quotient of $GL(n)$ by P when $\mathcal{E} = \mathcal{O}_S^n$ in the sense of Mumford (see [GIT] 0.6 and [GME] 1.8.3).

Proof. We repeat a proof given in [EGA] I.9.7. By the argument preceding the theorem, we may assume that $\mathcal{E} = \mathcal{O}_S^n$. We choose a base $I = \{v_1, v_2, \ldots, v_n\}$ of $\mathcal{O}_S^n$ over $\mathcal{O}_S$ and consider a subset J of order m of I. We then define a subset $G_J(S') \subset \underline{\mathrm{Grass}}_{\mathcal{E},m}(S')$ by

$$G_J(S') = \left[\phi : \mathcal{E}_{S'} \twoheadrightarrow \mathcal{F} \middle| \phi(\sum_{v\in J} \mathcal{O}_{S'} v) = \mathcal{F}\right] \subset \underline{\mathrm{Grass}}_{\mathcal{E},m}(S'),$$

where $\mathcal{E}_{S'} = f^*\mathcal{E}$ for $f : S' \to S$. By definition, $\underline{\mathrm{Grass}}_{\mathcal{E},m}(S') = \bigcup_J G_J(S')$ for all S-schemes S'. If $S'' \xrightarrow{g} S'$ is an S-morphism and $(\phi : \mathcal{E}_{S'} \twoheadrightarrow \mathcal{F}) \in G_J(S')$, by the right-exactness of g^*, we find that $g^*(\phi)(\sum_{v\in J} \mathcal{O}_{S''} v) = g^*\mathcal{F}$, and hence $G_J : SCH_{/S} \to SCH$ is a subfunctor of $\underline{\mathrm{Grass}}_{\mathcal{E},m}$.

We now show that G_J is represented by an affine scheme $\mathcal{G}_J$. Let $\mathcal{E}_J = \sum_{v\in J} \mathcal{O}_{S'} v$. Then $\phi : \mathcal{E}_{J/S'} \to \mathcal{F}$ is a surjective homomorphism of a locally free sheaf of the same rank; so, it is an isomorphism. Let $X = \mathbf{V}(\mathcal{E} \otimes \mathcal{O}_S^m)$ representing $S' \mapsto \mathrm{Hom}_{\mathcal{O}_{S'}}(\mathcal{E}_{S'}, \mathcal{O}_{S'}^m)$ and $Y = \mathbf{V}(\mathcal{O}_S^m \otimes \mathcal{O}_S^m)$ representing $S' \mapsto \mathrm{End}_{\mathcal{O}_{S'}}(\mathcal{O}_{S'}^m)$ (see (6.2)). Define the morphism of functors $\alpha_{S'}, \beta_{S'} : \underline{X}(S') \to \underline{Y}(S')$ by $\alpha_{S'}(v : \mathcal{E}_{S'} \to \mathcal{O}_{S'}^m) = v \circ \iota_{J,S'}$ and $\beta_{S'}(v) = \mathrm{id}_{\mathcal{O}_{S'}^m}$, where $\iota_{J,S'} : \mathcal{O}_{S'}^J = \mathcal{O}_{S'}^m \to \mathcal{E}_{S'}$ is given by $\iota_{J,S'}((a_v)_{v\in J}) = \sum_{v\in J} a_v v$. Define $\mathcal{G}_J$ by the following fibered product,

$$\begin{array}{ccc} \mathcal{G}_J = X \times_Y X & \longrightarrow & X \\ \downarrow & & \downarrow \alpha \\ X & \xrightarrow[\beta]{} & Y. \end{array}$$

Then we see

$$\underline{\mathcal{G}}_J(S') = \left\{(\phi, w) \in \mathrm{Hom}_{\mathcal{O}_{S'}}(\mathcal{E}_{S'}, \mathcal{O}_{S'}^m)^2 \middle| \phi \circ \iota_{J,S'} = \mathrm{id}_{\mathcal{O}_{S'}^J}\right\}$$
$$\cong \left\{\phi : \mathcal{E}_{S'} \to \mathcal{O}_{S'}^m \middle| \phi(\sum_{v\in J} \mathcal{O}_{S'} v) = \mathcal{O}_{S'}^m\right\} = G_J(S').$$

Thus the scheme $\mathcal{G}_J$ represents the functor G_J. Since X and Y are affine over S and X is affine over Y, $\mathcal{G}_J$ is affine over S.

Take another subset $J' \subset I$ of order m. We show that $G_J \cap G_{J'}$ is represented by an open subscheme of $\mathcal{G}_J$. If $\phi \in G_J(S')$, then $v = \phi \circ \iota_{J,S'} : \mathcal{O}_{S'}^m \to \mathcal{F}$ is an isomorphism. Write its inverse as $w : \mathcal{F} \to \mathcal{O}_{S'}^m$. Then $w \circ \iota_{S',J}$ is an

element in $\mathrm{End}_{\mathcal{O}_{S'}}(\mathcal{O}_{S'}^m)$; so, we have $\det(w \circ \iota_{S',J'}) \in \mathbf{G}_a(S')$. The association $\phi \mapsto \det(w \circ \iota_{S',J'})$ induces a functorial map of G_J into $\underline{\mathbf{G}}_a$; so, by Yoneda's lemma ([GME] Lemma 1.4.1), we have a morphism of S-schemes $\delta : \mathcal{G}_J \to \mathbf{G}_a$. Regarding δ as a section of $\mathcal{O}_{\mathcal{G}_J}$, we find that $G_J \cap G_{J'}(S') = \underline{\mathcal{G}}_J\left[\frac{1}{\delta}\right](S')$. This shows that the intersection $G_J \cap G_{J'}$ is represented by an open affine subscheme of $\mathcal{G}_J$ (and hence of $\mathcal{G}_{J'}$), and $\{\mathcal{G}_J\}_J$ glue into a scheme $\mathrm{Grass}_{\mathcal{E},m}$ representing the functor $\underline{\mathrm{Grass}}_{\mathcal{E},m}$. For each valuation ring R with $f : Spec(R) \to S$ and quotient field $K = Frac(R)$, $(f_K^*\mathcal{E} \xrightarrow{\phi} \mathcal{F}) \in \underline{\mathrm{Grass}}_{\mathcal{E},m}(K)$, we can extend $\mathcal{F} = \widetilde{V}$ uniquely to $Spec(R)$ as follows. Regarding $f_K : K^n \to V$, we have $M = R^n/(\mathrm{Ker}(f_K) \cap R^n)$. Then defining $\overline{\mathcal{F}} = \widetilde{M}$, we find an extension $f : \widetilde{R}^n \to \overline{\mathcal{F}}$. Plainly this is a unique torsion-free extension of $\mathcal{F}$. Thus by the valuative criterion of properness ([ALG] II.4.7 and [GME] Theorem 1.9.2), $\mathrm{Grass}_{\mathcal{E},m}$ is proper over S.

We give a sketch of a proof of the projectivity of the scheme $\mathrm{Grass}_{\mathcal{E},m}$. Since the property is local on S, we may assume that $S' = Spec(R)$ and $S = Spec(\mathcal{B})$ for an algebra $\mathcal{B}$. Then we write $\mathcal{E} = \widetilde{R}^n$. For each surjection $R^n \twoheadrightarrow M$ for a flat R-module M of rank m, further shrinking S', we may assume that $M \cong R^m$. Having a surjection $\phi : R^n \twoheadrightarrow R^m$ is equivalent to having an m-subset J of a base I of R^m such that $\iota_{S',J} : R^m \hookrightarrow R^m$ composed with ϕ is an isomorphism of R-modules. This in turn is equivalent to $\det(\phi \circ \iota_{S',J}) \in R^\times$. This shows that $(\mathcal{E}_{S'} \xrightarrow{\phi} \mathcal{F}) \mapsto ((\bigwedge^m \mathcal{E})_{S'} \xrightarrow{\wedge^m \phi} \bigwedge^m \mathcal{F})$ gives rise to a functorial injection $\iota : \underline{\mathrm{Grass}}_{\mathcal{E},m}(S') \hookrightarrow \underline{\mathrm{Grass}}_{(\bigwedge^m \mathcal{E}),1}(S')$, where $r+1 = \mathrm{rank}_{\mathcal{O}_S} \bigwedge^m \mathcal{E}$. The point of the projective space associated with a point $\mathcal{E}_{S'} \xrightarrow{\phi} \mathcal{F}$ in the Grassmannian is called the *Plücker coordinate* of $\mathcal{E}_{S'} \xrightarrow{\phi} \mathcal{F}$. Thus the morphism is an immersion. Since $\mathrm{Grass}_{\mathcal{E},m}$ is proper over S, it has to be a closed immersion (cf. [GME] 1.9.1).

As we remarked just after stating the theorem, even if $\mathcal{E}$ is merely locally free, $\mathrm{Grass}_{(\bigwedge^m \mathcal{E}),1}$ is projective, although $\mathrm{Grass}_{(\bigwedge^m \mathcal{E}),1}$ itself may not be a projective space. This shows the Plücker coordinate embedding: $\mathrm{Grass}_{\mathcal{E},m} \hookrightarrow \mathrm{Grass}_{(\bigwedge^m \mathcal{E}),1}$ composed with a projective embedding of $\mathrm{Grass}_{(\bigwedge^m \mathcal{E}),1}$ gives a projective embedding of $\mathrm{Grass}_{\mathcal{E},m}$. □

6.1.3 Flag Varieties

We can further generalize our construction of Grassmannians to flag varieties. We follow [EGA] I.9.9. We consider the following functor from S-SCH to $SETS$,

$$\underline{\mathrm{Flag}}_B(S') = \left[\pi_j : \mathcal{E}_{S'} \twoheadrightarrow \mathcal{E}_j \middle| \begin{array}{l}\mathrm{Ker}(\pi_{j+1}) \subset \mathrm{Ker}(\pi_j) \text{ and} \\ \mathcal{E}_j \text{ is locally free of rank } n-j \ (1 \le j \le n-1)\end{array}\right].$$

Here the subscript B indicates a split Borel subgroup of $GL(n)$, since $\underline{\mathrm{Flag}}$ is represented by a projective scheme $\mathrm{Flag}_B = GL(n)/B$ if $\mathcal{E} = \mathcal{O}_S^n$ as we show below (Theorem 6.2). Since $\mathcal{E} \cong \mathcal{O}_S^n$ locally, covering S by open subschemes

O over which $\mathcal{E}$ is free, we can glue schemes over the open Os representing $\underline{\mathrm{Flag}}_{B/O}$ to get a scheme representing $\underline{\mathrm{Flag}}$ over S. The association $(\pi_j : \mathcal{E}_{S'} \to \mathcal{E}_j) \mapsto (\pi_j)_j$ induces a functorial map,

$$\iota_{S'} : \underline{\mathrm{Flag}}_B(S') \hookrightarrow \underline{\mathrm{Grass}}_{\mathcal{E},1}(S') \times_S \underline{\mathrm{Grass}}_{\mathcal{E},2}(S') \times_S \cdots \times_S \underline{\mathrm{Grass}}_{\mathcal{E},n-1}(S'),$$

which is an injection for all j. Thus if $\underline{\mathrm{Flag}}_B$ is representable by a scheme $\mathrm{Flag}_{B/S}$, it is an immersion. As in the case of $\mathrm{Grass}_{\mathcal{E},j}$, we can show that $\mathrm{Flag}_{B/S}$ is proper; so, the morphism ι is a closed immersion. Thus $\mathrm{Flag}_{B/S}$ is projective if it is representable ([EGA] I.9.9.3).

Theorem 6.2 *The functor* $\underline{\mathrm{Flag}}_{B/S}$ *for a locally free sheaf* $\mathcal{E}$ *of rank* n *over* S *is representable by a projective scheme* $\mathrm{Flag}_{B/S}$.

Proof. By the argument preceding the theorem, we only need to prove the representability assuming $\mathcal{E} = \mathcal{O}_S^n$. Write

$$V = \mathrm{Grass}_{\mathcal{E},1} \times_S \mathrm{Grass}_{\mathcal{E},2} \times_S \cdots \times_S \mathrm{Grass}_{\mathcal{E},n-1}\,.$$

We have a universal morphism $\boldsymbol{\pi}_j : \mathcal{E}_V \to \mathbf{E}_j$. Let $\mathbf{M} = \mathrm{Ker}(\boldsymbol{\pi}_j \oplus \boldsymbol{\pi}_{j+1})$. Then for each $x = (\pi_j, \pi_{j+1}) \in V$, $\dim \mathbf{M} \otimes_{\mathcal{O}_Y} k(x) = n - j - 1$ implies that $\mathrm{Ker}(\pi_j) \supset \mathrm{Ker}(\pi_{j+1})$, otherwise $\dim \mathbf{M} \otimes_{\mathcal{O}_Y} k(x) \leq n - j - 2$. Then

$$V_j = \{x \in V \,|\, \dim \mathbf{M} \otimes_{\mathcal{O}_Y} k(x) = n - j - 1\}$$

is a closed subscheme of the support of $\mathbf{M}$ (cf. [GME] Lemma 4.1.3). Thus $\mathrm{Im}(\iota) = \bigcap_j V_j$ is a closed subscheme of V. This shows the desired representability of $\underline{\mathrm{Flag}}_{B/S}$. □

We can generalize the construction of flag varieties to vector bundles over Flag_B. We consider the following functor,

$$\underline{\mathrm{Flag}}_U(S') = \left[(\pi_j, \phi_j)\middle|(\pi_j) \in \mathrm{Flag}_B(S') \text{ and } \phi_j : \mathrm{Ker}(\pi_{j-1})/\mathrm{Ker}(\pi_j) \cong \mathcal{O}_{S'}\right].$$

Here we understand that $\mathrm{Ker}(\pi_0) = \mathcal{E}$, and j runs over all integers between 0 and $n-1$. If $\mathcal{E} \cong \mathcal{O}_S^n$ and S is affine, writing $\mathbf{1} = (\pi_j, \phi_j)$ for the standard flag $\pi_j : \mathcal{O}_S^n \to \mathcal{O}_S^{n-j}$, projecting column vectors down to lower $n-j$ coordinates, the upper unipotent subgroup U of $GL(n) = \mathrm{Aut}_{\mathcal{O}_S}(\mathcal{O}_S^n)$ is the stabilizer of $\mathbf{1}$. Therefore $\underline{\mathrm{Flag}}_U$ is represented by $\mathrm{Flag}_{U/S} = GL(n)/U$. In general, cover S by open affine schemes $S_i = Spec(\mathcal{B}_i)$ so that we have an isomorphism $\mathcal{E}|_{S_i} \cong \mathcal{O}_{S_i}^n$. On $S_{ij} = S_i \cap S_j$, by the universality, $\mathrm{Flag}_{U/S_i} \times_S S_{ij}$ is canonically isomorphic to $\mathrm{Flag}_{U/S_j} \times_S S_{ij}$. Thus these schemes glue each other, giving rise to a scheme $\mathrm{Flag}_{U/S}$ representing $\underline{\mathrm{Flag}}_U$. Obviously $\mathrm{Flag}_{U/S}$ is a T-torsor for the maximal split torus $T \subset GL(n)$. Here the action of T on Flag_U is given by $(\pi_j, \phi_j) \mapsto (\pi_j, t_j\phi_j)$ for $(t_1, \ldots, t_n) \in T = \overbrace{\mathbb{G}_m \times \cdots \times \mathbb{G}_m}^{n}$.

Let $\pi : \mathrm{Flag}_U \to \mathrm{Flag}_B$ be the projection $(\pi_j, \phi_j) \mapsto (\pi_j)$. Then for a character $\kappa \in X(T)$, we define a sheaf $\mathcal{E}^\kappa(O) = H^0(\pi^{-1}(O), \mathcal{O}_{\mathrm{Flag}_U}[\kappa])$ for each open subset $O \subset \mathrm{Flag}_B$. Then $\mathcal{E}^\kappa$ is a locally free sheaf on Flag_B. Since $f : \mathrm{Flag}_B \to S$ is proper flat over S, we find that $f_*\mathcal{E}^\kappa$ (which we again write $\mathcal{E}^\kappa$) is a locally free sheaf on S. In this way, we can associate a κ-power $\mathcal{E}^\kappa$ of the original locally free sheaf $\mathcal{E}$ for each weight $\kappa \in X(T)$ with respect to (B, T), which is non-zero if and only if κ is a dominant weight of $GL(n)$ with respect to the pair (B, T).

6.1.4 Flat Quotient Modules

Let $f : X \to S$ be a projective scheme over a (separated) Noetherian connected scheme S of relative dimension n. Here the word " projective" means that we have a closed immersion $\iota : X_{/S} \hookrightarrow \mathbf{P}^N_{/S}$. Thus X has a very ample invertible sheaf $\mathcal{O}_X(1) = \iota^*\mathcal{O}_{\mathbf{P}^N}(1)$. The sheaf of graded algebras $\mathcal{A} = \bigoplus_{n=0}^\infty f_*(\mathcal{O}_X(1)^n)$ determines X as $X = Proj_S(\mathcal{A})$.

For a coherent sheaf $\mathcal{M}$ on X, we write $\mathcal{M}(k)$ for $\mathcal{M} \otimes_{\mathcal{O}_X} \mathcal{O}_X(k)$ and define a sheaf of graded modules $M = \bigoplus_{k\geq 0} M_k$ by $M_k = f_*\mathcal{M}(k)$. Then M is a graded $\mathcal{A}$-module of finite type, and we have $\mathcal{M} = \widetilde{M}$, where $\widetilde{M}$ is the sheaf on $Proj_S(\mathcal{A})$ associated with the graded module M (e.g., [ALG] II.7 or [GME] 1.3.3). First removing finitely many graded pieces of M does not alter $\mathcal{M} = \widetilde{M}$. If $\mathcal{M}$ is $\mathcal{O}_S$-*flat*, for each geometric point $s = Spec(k(s)) \in S$, there is a polynomial $P_{\mathcal{M}}(X)$ such that

$$\chi(\mathcal{M}(n)) = \sum_{j=0}^{\dim X(s)} (-1)^j \dim_{k(s)} H^j(X(s), \mathcal{M}(n) \otimes_{\mathcal{O}_S} k(s)) = P_{\mathcal{M}}(n),$$

since the Euler characteristic is additive with respect to exact sequences. For sufficiently large n, the ampleness of $\mathcal{O}(1)$ tells us (a theorem of Serre: [ALG] III.5.2) that $H^j(X(s), \mathcal{M}(n) \otimes_{\mathcal{O}_S} k(s)) = 0$ if $j > 0$. Thus $P_{\mathcal{M}}(n)$ gives the dimension of $H^0(X(s), \mathcal{M}(n) \otimes_{\mathcal{O}_S} k(s))$, which is equal to the $\mathcal{O}_S$-rank of $f_*\mathcal{M}$ (by flatness of $\mathcal{M}$); so, $P_{\mathcal{M}}$ is independent of the choice of $s \in S$ (connectedness of S). When $S = Spec(\mathcal{B})$, M is associated with a graded $\mathcal{B}$-module, and $P_{\mathcal{M}}(n)$ is the Hilbert polynomial of this graded module (cf. [CRT] Section 13).

For a given coherent sheaf $\mathcal{F}_{/X}$, we consider the following contravariant functor defined over locally Noetherian S-schemes $S' \xrightarrow{\phi} S$ (inducing $\phi_X : X_{S'} = X \times_S S' \to X$),

$$\underline{\mathrm{Quot}}_{\mathcal{F}/X/S}(S') = \left[\pi : \phi_X^*\mathcal{F} \twoheadrightarrow \mathcal{M} \middle| \mathcal{M} \text{ is a coherent } \mathcal{O}_{X_{S'}}\text{-module flat over } \mathcal{O}_{S'}\right].$$

The isomorphism between the πs are defined in the same manner as in the case of the functor $\underline{\mathrm{Grass}}$. Here we do not assume that $\mathcal{F}$ is $\mathcal{O}_S$-flat.

For simplicity, we always assume that schemes S' (and also S) are Noetherian. Each $\mathcal{M} \in \underline{\mathrm{Quot}}_{\mathcal{F}/X/S}(S')$ has its Hilbert polynomial $P_{\mathcal{M}}$, and obviously for $g : S'' \to S'$, $g^*\mathcal{M}$ has the same Hilbert polynomial. Thus we can split the functor as

$$\underline{\mathrm{Quot}}_{\mathcal{F}/X/S} = \bigsqcup_P \underline{\mathrm{Quot}}^P_{\mathcal{F}/X/S},$$

where $\underline{\mathrm{Quot}}^P_{\mathcal{F}/X/S}(S') = \left[\pi : \phi_X^*\mathcal{F} \twoheadrightarrow \mathcal{M} \in \underline{\mathrm{Quot}}_{\mathcal{F}/X/S}(S') \middle| P_{\mathcal{M}} = P\right]$. Here is a theorem of Grothendieck.

Theorem 6.3 *Let the notation be as above. Suppose that S is Noetherian, that $\mathcal{F}_{/S}$ is coherent, and that $X_{/S}$ is a projective scheme of finite type with $\dim_S X = n$. Then $\underline{\mathrm{Quot}}^P_{\mathcal{F}/X/S}$ is representable by a projective scheme $Q^P = \mathrm{Quot}^P_{\mathcal{F}/X/S}$ of finite type over S. Thus we have, for any S-scheme S',*

$$\mathrm{Hom}_S(S', Q^P) \cong \left[\pi : f^*\mathcal{F} \twoheadrightarrow \mathcal{M} \in \underline{\mathrm{Quot}}_{\mathcal{F}/X/S}(S') \middle| P_{\mathcal{M}} = P\right].$$

We give a sketch of a proof of this theorem.

We recall $X = Proj_S(\mathcal{A})$ for a sheaf $\mathcal{A}$ of graded $\mathcal{O}_S$-algebras generated by $\mathcal{A}_1$. We cover $\underline{\mathrm{Quot}}^P_{\mathcal{F}/X/S}$ by subfunctors $\underline{Q}_j$ indexed by nonnegative integers j defined as follows. $\underline{Q}_j(S')$ consists of isomorphism classes of $\pi : \mathcal{F}_{/X_{S'}} \twoheadrightarrow \mathcal{M}_{/X_{S'}}$ satisfying the following three conditions:

(a) $R^i f_{S',*}\mathcal{M}(n)_{/X_{S'}} = 0$ for all $i > 0$ and $n \geq j$;
(b) $R^i f_{S',*}\mathcal{K}(n)_{/X_{S'}} = 0$ for all $i > 0$ and $n \geq j$, where $\mathcal{K} = \mathrm{Ker}(\pi)$;
(c) $\mathcal{A}_k f_{S',*}(\mathcal{K}(j)) = f_{S',*}(\mathcal{K}(j+k))$ for all $k > 0$.

Covering S' by affine schemes $Spec(B_i)$ and writing B_i as a union of Noetherian rings, we can reduce proofs to Noetherian S'; so, we assume that S' is Noetherian as we remarked already. By a theorem of Serre ([EGA] III Section 2 or [ALG] III.5.2), for any coherent sheaf $\mathcal{G}_{/X_{S'}}$, we have the vanishing: $R^i f_{S',*}\mathcal{G}(n)_{/X_{S'}} = 0$ for $n \gg 0$. Thus (a) and (b) will be satisfied for a given π for $j \gg 0$. Since $\mathcal{F}$ is coherent (and $X_{/S}$ is of finite type), it is finitely presented; so, $\mathcal{K}$ is finitely presented because $\mathcal{M}$ is finitely presented and locally free. Thus $\mathcal{K}$ is generated by $\mathcal{K}(j)$ for some j, and the last condition will be fulfilled if $j \gg 0$. Hence $\underline{\mathrm{Quot}}^P_{\mathcal{F}/X/S}(S')$ is covered by $\underline{Q}_j(S')$ for each S'.

Since the sequence $0 \to \mathcal{K} \to \mathcal{F} \to \mathcal{M} \to 0$ is exact, we have $\chi(\mathcal{F}) = \chi(\mathcal{K}) + \chi(\mathcal{M})$. Thus the conditions (a) and (b) tell us that Hilbert's polynomials $P_{\mathcal{K}}(n)$ and $P_{\mathcal{F}}(n)$ give the exact $\mathcal{O}_{S'}$-rank of $f_{S',*}\mathcal{K}$ and $f_{S',*}\mathcal{F}$. Vanishing of $R^1 f_{S',*}\mathcal{X}_{/X_{S'}} = 0$ implies that $f_{S',*}\phi_X^*\mathcal{X} = \phi^* f_*\mathcal{X}_{/X}$ ([EGA] III, 7.7.5, 7.7.10, 7.8.4, [ALG] III.12.10 or [GIT] 0.5); so, the conditions (a) and (b) are stable under base change. The tensor product is a right exact functor; so, the surjectivity of $p_k : \mathcal{A}_k \otimes f_{S',*}\mathcal{K}(j) \to f_{S',*}\mathcal{K}(j+k)$ is also kept under base change. Thus $\underline{Q}_j$ is a well-defined contravariant functor, and we have $\underline{\mathrm{Quot}}^P_{\mathcal{F}/X/S} = \bigcup_j \underline{Q}_j$. By the conditions (a–c), $\pi \in \underline{Q}_j(S')$ is determined

by $M_j = f_{S',*}\mathcal{M}(j)$ as a flat quotient of rank $P(j)$ of $F_j = f_{S',*}\mathcal{F}(j)$. Thus $\pi \mapsto (\pi_j : F_j \twoheadrightarrow M_j)$ induces a functorial injection $\underline{Q}_j(S') \hookrightarrow \underline{\mathrm{Grass}}_{F_j,P(j)}(S')$.

If π' is in the image of $\underline{Q}_j$, then π' satisfies:

(i) $F_{j+k}/\mathcal{A}_k K'$ for $K' = \mathrm{Ker}(\pi')$ is locally $\mathcal{O}_{S'}$-free of rank $P(j+k)$ for all $k \geq 0$. Here we consider $\mathcal{A}_k K'$ in $\mathcal{A} \cdot F$;

(ii) Define a graded module K'^* by $\mathcal{A} \cdot K'$. We require the associated sheaf $\mathcal{K}' = \widetilde{K'^*}_{/X_{S'}}$ on $X_{S'}$ to satisfy (b) and the quotient $\mathcal{M}' = \mathcal{F}/\mathcal{K}'$ to satisfy (a) (these (a) and (b) are open conditions: [ABV] Section 5 Corollary or [ALG] Theorem III.12.8).

For a graded $\mathcal{A}$-module $M = \bigoplus_k M_k$, putting $M(t) = \bigoplus_{k \geq t} M_k$, we have $\widetilde{M(t)} \cong \widetilde{M}$ as already remarked. Then the condition (i) ensures that $f_{S',*}\mathcal{M}'$ is locally $\mathcal{O}_{S'}$-free, and the image of $\underline{Q}_j(S')$ is characterized by (i) and (ii).

We prove the representability of $\underline{Q}_j$, assuming that $j = 0$. The general case follows from the same argument replacing 0 by j everywhere. Let $\pi^{univ} : F_{0/G} \twoheadrightarrow M_0$ ($F_{0/G} = F_0 \otimes_{\mathcal{O}_S} \mathcal{O}_G$) be the universal quotient defined over $G := \mathrm{Grass}$. Here we change our notation and write M_0 for the universal quotient of $F_{0/G}$ (with rank $P(0)$). Thus for any morphism $\pi' : F_0 \to M_0'$ over S' with M_0' locally free of rank $P(0)$, we have a unique morphism $\phi : S' \to G$ such that $\pi' = \phi^*\pi^{univ}$. Let $K = \mathrm{Ker}(\pi^{univ})$. Write $g : G \to S$ for the structure morphism. We consider the subset:

$$Z = \left\{ s \in G \middle| \begin{array}{l} \dim_{k(s)}(\mathcal{A}_k F_{0/G}/(\mathcal{A}_k K)) \otimes k(s) = P(k), \quad \text{and} \\ \text{the stalk } (\mathcal{A}_k F_{0/G}/(\mathcal{A}_k K))_s \text{ is } \mathcal{O}_{G,s}\text{-free for all } k \geq 0 \end{array} \right\}.$$

Write $M_k = \mathcal{A}_k F_{0/G}/(\mathcal{A}_k K)$ and put $M = \bigoplus_{k \geq 0} M_k$. Then $M_k = \mathcal{A}_k M_0$. The $\mathcal{O}_G$-module M is flat on a generic point of $Z_0 := Supp(M) = G$. Since flatness is an open condition, we find an open connected subscheme $V_0 \subset Z_0$ which is maximal among open subschemes V over which M is flat. Repeating this process, replacing M and Z_0 by $M \otimes_{\mathcal{O}_{Z_0}} \mathcal{O}_{Z_1}$ and $Z_1 = Z_0 - V_0$, we can split $Z_0 = \bigsqcup_i V_i$ into a finite disjoint union (as a topological space) of connected subschemes V_i so that $M \otimes_{Z_0} \mathcal{O}_{V_i}$ is flat over V_i. Then we find a polynomial $Q_i(n)$ such that $\mathrm{rank}_{\mathcal{O}_{V_i}}(M_n \otimes_{Z_0} \mathcal{O}_{V_i}) = Q_i(n)$ if $n \geq n_i$. By this fact, the subscheme

$$U_N = \left\{ s \in Z_0 \middle| \dim_{k(s)}(M_n \otimes_{\mathcal{O}_{Z_0}} k(s)) \leq P(n)\ 0 \leq \forall n \leq N \right\}$$

stabilizes as N grows. Since the function $s \mapsto \dim_{k(s)}(M_n \otimes_{\mathcal{O}_{Z_0}} k(s))$ is upper semicontinuous (e.g., [ABV] Section 5 Corollary or [ALG] III.12.8), the scheme U_N is an open subscheme of Z_0. We have now verified that on an open subscheme $U = U_\infty$ of Z_0, we have $\dim_{k(s)}(M_n \otimes_{\mathcal{O}_{Z_0}} k(s)) \leq P(n)$ (if $n \geq 0$) for all $s \in U$. We have an exact sequence $\mathcal{O}_U^q \xrightarrow{p_k} \mathcal{O}_U^{P(k)} \to M_k \to 0$, and Z is given by the closed subscheme of U on which all matrix coefficients of p_k vanish for all $k \geq 0$. Thus the image of Q_0 falls into $\underline{Z}$. The condition (ii) can be checked to be satisfied on an open subscheme of Z. Thus we have

Proposition 6.4 *The functor $\underline{Q}_J$ is represented by a quasi-projective scheme Q_J of finite type over S.*

Here the word "quasi-projective" means that the scheme has an open immersion into a projective scheme. Since Grass is projective, Q_J is quasi-projective.

The next step is to show that the increasing sequence of quasi-projective schemes $\{Q_J\}_J$ stabilizes after $j \geq N_0$; so, $\underline{\mathrm{Quot}}^P_{\mathcal{F}/X/S}$ is represented by a quasi-projective scheme. See [FGA] Exposé 221, Section 2 for an argument valid for a general $X_{/S}$.

We prove this fact only for $\mathcal{F} = \mathcal{O}_X$ in a manner slightly different from [FGA] Exposé 221. First assume X is the projective space $\mathbf{P}^n_{/S}$ and $\mathcal{F} = \mathcal{O}_X$. Writing $X = Proj(\mathcal{O}_S[t_0, \ldots, t_n])$ and $D_i \subset X$ for $Spec\left(\mathcal{O}_S\left[\frac{t_0}{t_i}, \ldots, \frac{t_n}{t_i}\right]\right)$, then $R^i f_* \mathcal{O}(j) = 0$ for all $i > 0$ if $j > n+1$ by a computation of cohomology groups by Čech cohomology with respect to the covering $X = \bigcup_{J=0}^n D_J$ (see [ALG] III.5). A version of the argument of Grothendieck for $X = \mathbf{P}^n$ to prove (a) and (b) for sufficiently large j and all $\mathcal{M}$ and $\mathcal{K}$ is as follows. Since $P_{\mathcal{F}} = P_{\mathcal{K}} + P_{\mathcal{M}}$ with $P_{\mathcal{M}} = P$, $P_{\mathcal{K}}$ is determined by P. Choosing homogeneous generators $x_1, \ldots, x_r$ of degree $-p$ of $\mathcal{K}$, we have a surjection $\pi_0 : \mathcal{O}(p)^r \to \mathcal{K}$ taking $(a_1, \ldots, a_r) \mapsto \sum_{i=1}^r a_i x_i$. Here r and p are determined by the first two leading terms of $P_{\mathcal{K}}$ and hence those of P. Let $\mathcal{K}_0 = \mathrm{Ker}(\pi_0)$. Then $rP_{\mathcal{O}(p)} = P_{\mathcal{K}_0} + P_{\mathcal{K}}$. Let $r_0 = r$ and $p = p_0$. The polynomial $P_{\mathcal{K}_0}$ is determined by $P_{\mathcal{K}}$. Thus the first two leading terms of $P_{\mathcal{K}_0}$ are bounded below and above independent of $\mathcal{K}$. Repeating this argument, we find an integer $N_1 \gg 0$ such that for integers $p_i > -N_1$ $(i = 0, 1, \ldots, n)$ we have the following exact sequences: $0 \to \mathcal{K}_i \to \mathcal{O}(p_i)^{r_i} \to \mathcal{K}_{i-1} \to 0$ with $\mathcal{K}_{-1} = \mathcal{K}$. By the associated long exact sequence, if $j \geq N_1 + n + 1$, $R^{q+1} f_* \mathcal{K}_i(j) \cong R^q f_* \mathcal{K}_{i-1}(j)$. Since cohomological dimension of $\mathbf{P}^n$ is n (as easily checked by Čech cohomology), for $i > 0$, $0 = R^{n+i+1} f_* \mathcal{K}_n(j) = R^{n+i} f_* \mathcal{K}_{n-1}(j) = \cdots = R^i f_* \mathcal{K}(j)$. By the same argument, $R^i f_* \mathcal{M}(j) = 0$ for all $i > 0$ and all $j \geq N_1 + n + 2 = N_0$. So Q_J stabilizes after $j \geq N_0$.

It is customary to identify $\pi \in \mathrm{Quot}_{\mathcal{O}_X/X/S}(S')$ with a closed immersion of $Spec_Q(\mathrm{Im}(\pi))$ into X; so, $\mathrm{Quot}^P_{\mathcal{O}_X/X/S}$ represents the contravariant functor

$$\underline{\mathrm{Hilb}}^P_{X/S}(S') = \{\text{closed subschemes of } X_{S'} \text{ flat over } S' \text{ with Hilbert polynomial } P\}.$$

This scheme is called the *Hilbert scheme* of X for the polynomial P. There is a direct way (as indicated in [GIT] 0.5.iii) of constructing $\mathrm{Hilb}^P_{X/S}$ for general $X \subset \mathbf{P}^n$ assuming its existence for $X = \mathbf{P}^n_{/S}$ as proved above. We give details of this. Let $X \subset \mathbf{P}^n_{/S}$ be a projective scheme. We write $H = \mathrm{Hilb}^P_{\mathbf{P}^n/S}$ and let $\pi : \mathcal{Z} \to H$ be the universal projective scheme over H with $i : \mathcal{Z} \hookrightarrow \mathbf{P}^n_{/H}$. In other words, for any closed subscheme $Z \subset \mathbf{P}^n_{/S'}$ for $S' \in SCH_{/S}$ with

Hilbert polynomial P, we have a unique S-morphism $\phi : S' \to H$ such that the following diagram commutes.

$$\begin{array}{ccc} \phi^*\mathcal{Z} = \mathcal{Z} \times_H S' & \xrightarrow{\subset} & \phi^*\mathbf{P}^n_{/H} = \mathbf{P}^n_{/H} \times_H S' \\ \wr\downarrow & & \wr\downarrow \\ Z & \xrightarrow[\subset]{} & \mathbf{P}^n_{/S'}. \end{array}$$

We define $X_H = X \times_{\mathbf{P}^n_{/S}} \mathbf{P}^n_{/H}$ and $\mathcal{Z}_X = \mathcal{Z} \times_{\mathbf{P}^n_{/H}} X_H = X \times_{\mathbf{P}^n_{/S}} \mathcal{Z}$ as a subscheme of $\mathbf{P}^n_{/H}$. We define $\mathrm{Hilb}^P_{X/S}$ to be the image of $\mathcal{Z}_X$ in H, that is, the topological image with reduced scheme structure. Since $\mathcal{Z}_X \to H$ is projective, the image $\mathrm{Hilb}^P_{X/S}$ is a well-defined closed subscheme of $\mathrm{Hilb}^P_{\mathbf{P}^n/S}$. Then, if $Z_{/S'}$ is a closed subscheme of $X_{/S'}$, $\phi^*\mathcal{Z} = Z \subset X$, and hence $\phi^*(\mathcal{Z}_X) = Z$ with $\phi : S' \to H$ factoring through $\mathrm{Hilb}_{X/S}$. Thus we have

$$Z \subset X_{S'} = X \times_S S' \iff \phi \text{ factors through } \mathrm{Hilb}^P_{X/S}.$$

From this, $\mathcal{Z}_X \to \mathrm{Hilb}^P_{X/S}$ is the universal closed subscheme in $X \times_S \mathrm{Hilb}^P_{X/S}$, and $\mathrm{Hilb}_{X/S}$ represents the functor $\underline{\mathrm{Hilb}}^P_{X/S}$.

We now finish the proof of the following theorem ([FGA] Exposé 221, 3.2).

Theorem 6.5 (Grothendieck) *For a projective scheme $X_{/S}$ over a Noetherian connected scheme S and a numerical polynomial $P(t) \in \mathbb{Q}[t]$, the functor $\underline{\mathrm{Hilb}}^P_{X/S}$ is represented by a projective scheme $\mathrm{Hilb}^P_{X/S}$.*

Proof. We only need to prove projectivity of $\mathrm{Hilb}^P_{X/S}$ by the valuative criterion. Let $\pi : \mathcal{O}_{X_\eta} \to \mathcal{M}_{/\eta} \in \underline{\mathrm{Quot}}^P_{\mathcal{O}_X/X/S}(\eta)$ for $\eta = Spec(K)$ of the field K of fractions of a discrete valuation ring V. Then we define $\mathrm{Ker}(\pi)_{/S}$ for $S' = Spec(V)$ by the largest subsheaf over S' of $\mathcal{O}_{X_{S'}}$ inducing $\mathrm{Ker}(\pi)$, that is, $\mathcal{O}_{X_{S'}} \cap \mathrm{Ker}(\pi)$, which is a coherent sheaf with quotient $\mathcal{M}_{/S'}$ locally free over $X_{S'}$ inducing $\mathcal{M}_{/\eta}$ after tensoring K, because V is a discrete valuation ring. Thus the point $\pi \in \underline{\mathrm{Quot}}^P_{\mathcal{O}_X/X/S}(\eta)$ extends to $\underline{\mathrm{Quot}}^P_{\mathcal{O}_X/X/S}(S')$. Since $\mathrm{Quot}^P_{\mathcal{O}_X/X/S} = \bigcup_{j=0}^{N_0} Q_j$ is quasi-projective, it is separated; so, it is proper. Since $\mathrm{Quot}^P_{\mathcal{O}_X/X/S}$ is quasi-projective, it has to be projective. □

6.1.5 Morphisms Between Schemes

Let $Y_{/S} \xrightarrow{f} X_{/S}$ be a morphism of S-schemes. Let $\underline{\mathrm{Sec}}_{Y/X/S} : SCH_{/S} \to SETS$ be contravariant functors given by

$$\underline{\mathrm{Sec}}_{Y/X/S}(S') = \mathrm{Hom}_{X_{S'}}(X_{S'}, Y_{S'}).$$

Each section $s : X_{S'} \to Y_{S'}$ defines a closed subscheme $\sigma(X_{S'})$ of $Y_{S'}$ isomorphic to $X_{S'}$ via $f_{S'}$. Thus if X is flat over S, $\sigma(X_{S'})$ is flat over S'; so,

$\sigma(X_{S'}) \in \underline{\mathrm{Hilb}}_{Y/S}(S')$. Write $H = \mathrm{Hilb}_{Y/S} = \bigsqcup_P \mathrm{Hilb}^P_{Y/S}$. Then we have the universal closed subscheme Z of $Y_H = Y \times_S H$ satisfying the commutative diagram

$$\begin{array}{ccc} Z & \xrightarrow[\subset]{i} & Y_H \\ \downarrow & & \downarrow \\ H & = & H \end{array}$$

such that for any S-scheme S' and a closed subscheme $V \hookrightarrow Y_{S'}$ flat over S', we have a unique morphism $\phi_V : S' \to H$ over S such that the pullback of the above square by ϕ_V is identical to

$$\begin{array}{ccc} V & \xrightarrow{\subset} & Y_{S'} \\ \downarrow & & \downarrow \\ S' & = & S'. \end{array}$$

We consider the S-subschemes $U \subset H$ such that $\widetilde{f}_H : Z \subset Y_H \xrightarrow{f_U} X_H$ for a given $f : Y \to X$ induces an isomorphism $\widetilde{f}_U : Z_U \cong X_U$. If there exists the largest $\widetilde{U}$ with this property, $\widetilde{U}$ represents the functor $\underline{\mathrm{Sec}}_{Y/X/S}$. The largest $\widetilde{U}$ is given by $\widetilde{U} = Z - \left(Supp(\mathrm{Coker}(\widetilde{f}^*_H : \widetilde{f}^*_H \mathcal{O}_{X_H} \to \mathcal{O}_Z)) \cup Supp(\mathrm{Ker}(\widetilde{f}^*_H))\right)$. Thus $\widetilde{U}$ is an open subscheme of H; so, each connected component of $\widetilde{U}$ is quasi-projective, since $\mathrm{Hilb}^P_{Y/S}$ is projective over S.

We consider two S-schemes X and Y. Then $Y_X = Y \times_S X$ has projection $p_X : Y_X \to X$. Note that $\underline{\mathrm{Sec}}_{Y_X/X/S}(S') = \mathrm{Hom}_{S'}(X_{S'}, Y_{S'})$. Thus if $X_{/S}$ is flat, we can apply the above argument to $p_X : Y_X \to X$ and the functor $\underline{\mathrm{Hom}}_S(X, Y) : S' \mapsto \mathrm{Hom}_{S'}(X_{S'}, Y_{S'})$ is representable.

Theorem 6.6 *Let $X_{/S}$ and $Y_{/S}$ be projective schemes over a Noetherian scheme S. Suppose $X_{/S}$ is flat. Then the functors $\underline{\mathrm{Sec}}_{Y/X/S}$ for an S-morphism $Y \xrightarrow{f} X$ and $\underline{\mathrm{Hom}}_S(X, Y)$ are representable by S-schemes $\mathrm{Sec}_{Y/X/S}$ and $H(X, Y)_{/S}$ over S, respectively. Each connected component of $\mathrm{Sec}_{Y/X/S}$ and $H(X, Y)_{/S}$ is quasi-projective over S.*

By construction, the scheme representing these functors may not be of finite type over S, because $\mathrm{Hilb}_{X/S}$ could have infinitely many components. However, each connected component of the scheme is of finite type over S.

If sections $s : S \hookrightarrow X$ and $s' : S \hookrightarrow Y$ are given, we may consider a subfunctor

$$\underline{\mathrm{Hom}}^{s,s'}(X, Y)_{/S}(S') = \left\{\phi \in \mathrm{Hom}_{S'}(X_{S'/S'}, Y_{S'/S'}) \middle| \phi \circ s_{S'} = s'_{S'}\right\}.$$

Corollary 6.7 *Let the notation and the assumption be as in the theorem. The functor $\underline{\mathrm{Hom}}^{s,s'}(X, Y)_{/S}$ is representable by an S-scheme $H^{s,s'}(X, Y)_{/S}$ whose connected components are quasi-projective over S.*

Proof. Consider the functorial map $\mathrm{Hom}_{S'}(X_{S'}, Y_{S'}) \ni \phi \mapsto \phi(s_{S'}) \in \underline{Y}(S')$ which induces a morphism $\sigma : \underline{\mathrm{Hom}}(X,Y)_{/S} \to Y$ by Yoneda's lemma (see [GME] Lemma 1.4.1). Then the functor $\underline{\mathrm{Hom}}^{s,s'}(X,Y)_{/S}$ is again representable by a scheme $H^{s,s'}(X,Y)_{/S} = H(X,Y)_{/S} \times_{Y,\sigma,s'} S$ over S. □

Corollary 6.8 *Let the notation and the assumption be as in the theorem. For a section $s : S \hookrightarrow X$, consider the following functor,*

$$\underline{E}^s_{X/S}(S') = \left\{\phi \in \mathrm{End}_{S'}(X_{S'/S'}) \middle| \phi \circ s_{S'} = s_{S'}\right\}.$$

The functor $SCH_{/S} \ni S' \mapsto \underline{E}^s_{S'}(X_{S'}) \in SETS$ is representable by a scheme $E^s_{X/S}$ over S whose connected components are quasi-projective over S.

6.1.6 Abelian Schemes

An *abelian scheme* $A_{/S}$ is a smooth geometrically connected group scheme proper over a separated locally Noetherian base S. An abelian scheme $A_{/S}$ with a polarization is projective (existence of polarization is equivalent to existence of an ample line bundle; see [ABV] Sections 16–17). Hereafter we always deal with polarized abelian schemes (so all abelian schemes we study are projective).

We can drop the "local Noetherian" hypothesis, because a smooth geometrically connected and proper group scheme over any base is a base change of such a scheme over a locally Noetherian base (cf. [DAV] I.1.2).

For simplicity, we actually suppose that S is a Noetherian scheme. Since A is a group, it has the identity section $\mathbf{0} : S \to A$. As in the elliptic curve case, any S-morphism $\phi : A \to A'$ of abelian schemes is a homomorphism if $\phi(\mathbf{0}_A) = \phi \circ \mathbf{0} = \mathbf{0}_{A'}$ (by the Rigidity lemma: [ABV] Section 4, [GIT] 6.4 and [GME] 4.1.5). In particular, if A is an abelian scheme over S, every scheme endomorphism of $A_{/S}$ keeping the zero section is a homomorphism of the group structure. Thus $E^{\mathbf{0}}_{A/S}$ is a ring scheme associated with the functor $S' \mapsto \underline{E}^{\mathbf{0}}_{S'}(A_{S'})$ with values in the category of rings.

Assume that A is an abelian scheme over a Noetherian base S. Take a connected component $E \subset E^{\mathbf{0}}_{A/S}$. Each connected component of $E^{\mathbf{0}}_{A/S}$, in particular E, is quasi-projective over S. Since S is Noetherian, E is of finite type over S, because of our construction $E^{\mathbf{0}}_{A/S} = H(A,A)_{/S} \times_{A,\sigma,\mathbf{0}} S$.

Suppose we have a discrete valuation ring $\mathcal{V}$ with field of fractions K and a morphism $\eta : Spec(K) \to E$ which is over a morphism $i : Spec(\mathcal{V}) \hookrightarrow S$. Then η gives rise to a section of $\underline{E}^{\mathbf{0}}_{A/S}(K)$. Since homomorphisms of abelian schemes are kept under specialization (which we call the *rigidity of endomorphisms*; see [GME] Subsections 4.1.5–6 and [DAV] I.2.7), η extends to $Spec(\mathcal{V})$ uniquely. By the valuative criterion of properness, we find that E is projective over S. If ϕ is an endomorphism of the abelian scheme $A_{/S}$, $\mathrm{Ker}(\phi)$ is again a group scheme, because it is given by $A \times_{A,\phi,\mathbf{0}_S} S$. If $\dim_S \mathrm{Ker}(\phi) = 0$, $\mathrm{Ker}(\phi)$ is a

locally free group scheme of finite rank (see [GME] Theorem 4.1.17 (2)); in this case, we call ϕ an *isogeny*. Thus ϕ is finite and hence affine (cf. [ALG] Exercise II.5.17). This implies $\ker(\phi) = Spec(R)$ for a locally free $\mathcal{O}_S$-algebra R. We define the *degree* $\deg(\phi)$ of ϕ by the $\mathrm{rank}_{\mathcal{O}_S} R$. If $\dim_S \mathrm{Ker}(\phi) > 0$, we simply put $\deg(\phi) = 0$.

If the connected component $E \subset E^{\mathbf{0}}_{A/S}$ contains an isogeny, the degree is independent of the point of E. To see this, we consider $A_E = A \times_S E$. Then ϕ induces $\phi_E \in \mathrm{End}(A_{E/E})$, we have $\deg(\phi) = \deg(\phi_E)$, and $\deg(\phi_E)$ is constant on the connected component of E; so, it is constant on E. For an abelian variety over a field, the number of isogenies with a given positive degree is finite (positivity of the Rosati involution: see [ABV] Section 21). Thus E is projective and quasi-finite; so, E is finite over S ([GME] Proposition 1.9.11).

We can generalize this argument to a strongly quasi-projective semi-abelian scheme. A scheme $X \xrightarrow{f} S$ is *strongly projective* if $X_{/S}$ is finitely presented and there exists a locally free sheaf $\mathcal{E}_{/S}$ of a constant finite rank and we have a closed immersion $X \hookrightarrow \mathbf{P}(\mathcal{E})$ over S. If $X_{/S}$ has an open immersion into a strongly projective scheme $X'_{/S}$, we call $X_{/S}$ strongly quasi-projective. An abelian scheme $A_{/S}$ equipped with a polarization $\lambda : A \to {}^tA$ is strongly projective (over any base S), because for a locally ample line bundle L, $L_1 = L^{\otimes 2}$ is globally ample as remarked in Section 4.1.1 and $L_1^{\otimes 3}$ is very ample (e.g., [ABV] Section 17).

As before, we suppose that S is Noetherian and any scheme $X_{/S}$ we consider is of finite type over S. Then automatically $X_{/S}$ is finitely presented. If $X_{/S}$ is projective, then by definition (cf. [EGA] II.5.5), we have a coherent sheaf $\mathcal{E}_{/S}$ and a closed immersion of X into $\mathbf{P}(\mathcal{E})$. Covering $\mathcal{E}$ by $\mathcal{O}_S^{n+1}$, we have a closed immersion $\mathbf{P}(\mathcal{E}) \hookrightarrow \mathbf{P}^n_{/S}$. Thus $X_{/S}$ is strongly projective under our finiteness condition. Also if $X_{/S}$ is quasi-projective, it is strongly quasi-projective under our finiteness assumption. Thus the results stated in [NMD] Chapter 8 under some strong (quasi) projectivity assumptions hold under the corresponding projectivity assumptions in our setting.

A *semi-abelian scheme* $\mathcal{G}_{/S}$ is a smooth separated group scheme with geometrically connected fiber such that each geometric fiber is an extension of an abelian variety by a torus. The toric rank may depend on the fiber. If the abelian scheme has polarization, $\mathcal{G}_{/S}$ is quasi-projective. We assume that $\mathcal{G}_{/S}$ is a quasi-projective semi-abelian scheme which is an abelian scheme over a dense open subscheme O of S. The semi-abelian scheme over the toroidal compactification of the Mumford moduli extending the universal abelian scheme (see Theorem 6.20) constructed in [DAV] VI.1 is (strongly) quasi-projective (because its abelian variety quotient has polarization and the toroidal compactification is Noetherian). Then $\mathrm{Hilb}_{\mathcal{G}/S}$ exists as a quasi-projective scheme (we only proved this when $\mathcal{G}/S$ is projective; see [NMD] Theorem 8 in Section 8.2). Thus the scheme $E^{\mathbf{0}}_{\mathcal{G}/S}$ exists. It is known that any homomorphism of semi-abelian schemes $\mathcal{G}_O \to \mathcal{G}'_O$ defined over an open dense subscheme $O \subset S$ extends uniquely to $\mathcal{G}_{/S} \to \mathcal{G}'_{/S}$ (endomorphisms are kept under specializa-

tion; a theorem of M. Raynaud, [NMD] 7.4–5 and [DAV] I.2.7). Thus if $\mathcal{G}_{/S}$ is an abelian scheme over a dense open subscheme $O \subset S$, we have a unique extension of the scheme $E^{\mathbf{0}}_{\mathcal{G}_O/O}$ over O to a scheme $\mathrm{End}_{\mathcal{G}/S} \subset E^{\mathbf{0}}_{\mathcal{G}/S}$ over S which represents the functor $T \mapsto \mathrm{End}(\mathcal{G}_T/T)$ for the group endomorphism algebra $\mathrm{End}(\mathcal{G}_T/T)$. Applying the valuative criterion using this rigidity of endomorphisms, we find that $\mathrm{End}_{\mathcal{G}/S}$ has connected components each projective over S. This shows the following fact.

Corollary 6.9 *Let $\mathcal{G}_{/S}$ be a quasi-projective semi-abelian scheme over a Noetherian base S. Suppose that $\mathcal{G}$ is an abelian scheme over an open dense subscheme of S. Then the functor $T \mapsto \mathrm{End}_T(\mathcal{G}_{T/T})$ is represented by a scheme $\mathrm{End}_{\mathcal{G}/S} \subset E^{\mathbf{0}}_{\mathcal{G}/S}$ over S. Each connected component of $\mathrm{End}_{\mathcal{G}/S}$ is projective over S. If the connected component contains an isogeny, it is finite over S. Here $\mathrm{End}_T(\mathcal{G}_{T/T})$ denotes endomorphisms of $\mathcal{G}$ compatible with group structure on $\mathcal{G}$.*

This result enables us to prove

Corollary 6.10 *Let S be a Noetherian scheme, and $\mathcal{G}_{/S}$ be a quasi-projective semi-abelian scheme which is an abelian scheme over a dense open subscheme of S. Let D be a finite-dimensional simple algebra over $\mathbb{Q}$ and $O = O_D \subset D$ be a subalgebra such that O_D is of finite type over $\mathbb{Z}$ and $D = O_D \otimes_{\mathbb{Z}} \mathbb{Q}$. Then the functor $S'_{/S} \mapsto \left[(\mathcal{G}_{S'}, \theta \in \mathrm{Hom}_{alg}(O_D, \mathrm{End}(\mathcal{G}_{S'/S'})))\middle|\theta(1_D) = \mathrm{id}_{\mathcal{G}}\right]$ is representable by a scheme $E^{O}_{\mathcal{G}/S}$ over S, and each connected component of $E^{O}_{\mathcal{G}/S}$ is finite over S. If S is a scheme over a discrete valuation ring of residual characteristic $p > 0$, each connected component of $E^{O}_{\mathcal{G}/S}$ is unramified over S in the sense of* [SGA] I.3. *The algebra homomorphism $\theta : O_D \to \mathrm{End}(\mathcal{G}_{S'/S'})$ as above is in fact an embedding.*

Proof. We start with a general argument. Taking its connected component, we may assume that S is connected. Since $\mathrm{End}_{\mathcal{G}/S}$ is a disjoint union of projective schemes over S (by the above corollary), for any given connected projective scheme Y flat over S and each connected component $E_{/S}$ of $E^{\mathbf{0}}_{\mathcal{G}/S}$, the functor $T \mapsto \mathrm{Hom}_T(Y \times_S T, E \times_S T)$ is representable by a scheme $H(Y,E)_{/S}$ over S. Taking a disjoint union of $H(Y,E)$ indexed by connected components of $E^{\mathbf{0}}_{\mathcal{G}/S}$, the functor $T \mapsto \mathrm{Hom}_T(Y \times_S T, E^{\mathbf{0}}_{\mathcal{G}/S} \times_S T)$ is representable by a scheme $H(Y, E^{\mathbf{0}}_{\mathcal{G}/S})_{/S}$. If $\mathcal{Y}$ is a disjoint union of projective schemes Y flat over S (with a given S-point y), taking the fibered product of $H(Y, E^{\mathbf{0}}_{\mathcal{G}/S})_{/S}$ over S indexed by connected components Y of $\mathcal{Y}_{/S}$, again the functor $T \mapsto \underline{\mathrm{Hom}}^{y,1}_T(\mathcal{Y} \times_S T, E^{\mathbf{0}}_{\mathcal{G}/S} \times_S T)$ is representable by a scheme $H(\mathcal{Y}, E^{\mathbf{0}}_{\mathcal{G}/S})_{/S}$, where the superscript $\underline{\mathrm{Hom}}^{y,1}$ indicates that morphisms are required to take $y \in \mathcal{Y}(S)$ to the identity in $\mathrm{End}(\mathcal{G}_{/S})$ (see Corollary 6.7).

Now suppose that $\mathcal{Y}$ is a ring scheme flat over S with the identity section $\mathbf{1}_{\mathcal{Y}}$. Then the compatibility of $\phi \in H(\mathcal{Y}, E^{\mathbf{0}}_{\mathcal{G}/S})(T)$ with multiplication and addition of the two ring schemes $\mathcal{Y}$ and $E^{\mathbf{0}}_{\mathcal{G}/S}$ is a closed condition. For

example, the condition $\phi \circ +_{\mathcal{Y}} = + \circ (\phi \times \phi)$ for the addition $+_{\mathcal{Y}}$ for $\mathcal{Y}$ and $+$ for $E^{\mathbf{0}}_{\mathcal{G}/S}$ can be interpreted as the identity $+_* \circ \Delta = +^*_{\mathcal{Y}}$ for the diagonal map $\Delta : H(\mathcal{Y}, E^{\mathbf{0}}_{\mathcal{G}/S}) \to H(\mathcal{Y}, E^{\mathbf{0}}_{\mathcal{G}/S}) \times_S H(\mathcal{Y}, E^{\mathbf{0}}_{\mathcal{G}/S})$, where $+_* : H(\mathcal{Y}, E^{\mathbf{0}}_{\mathcal{G}/S}) \times_S H(\mathcal{Y}, E^{\mathbf{0}}_{\mathcal{G}/S}) \to H(\mathcal{Y} \times_S \mathcal{Y}, E^{\mathbf{0}}_{\mathcal{G}/S})$ is $(\phi \times \phi') \mapsto (+ \circ (\phi \times \phi'))$ and $+^*_{\mathcal{Y}} : H(\mathcal{Y}, E^{\mathbf{0}}_{\mathcal{G}/S}) \to Z$ is given by $\phi \mapsto (\phi \circ +_{\mathcal{Y}})$. If we have two morphisms of S-schemes $a, b : T \rightrightarrows T'$, the maximal subscheme $T_{a=b} \subset T$ on which $a = b$ is given by $\Delta_T^{-1}(T \times_{T'} T)$ for the diagonal map $\Delta_T : T \to T \times_{T'} T$, where the fibered product is taken with respect to a and b. Applying this to $T = H(\mathcal{Y}, E^{\mathbf{0}}_{\mathcal{G}/S})$ and $T' = H(\mathcal{Y} \times_S \mathcal{Y}, E^{\mathbf{0}}_{\mathcal{G}/S})$ and $(a, b) = (+_* \circ \Delta, +^*_{\mathcal{Y}})$, the functor: $T \mapsto \{\text{additive morphisms in } \mathrm{Hom}_S(\mathcal{Y}, E^{\mathbf{0}}_{\mathcal{G}/S})\}$ is represented by the closed subscheme Z of $H(\mathcal{Y}, E^{\mathbf{0}}_{\mathcal{G}/S})$. By the same argument applied to Z and the multiplication of the two ring schemes and taking the intersection with $H^{1_{\mathcal{Y}}, \mathrm{id}_{\mathcal{G}}}(\mathcal{Y}, E^{\mathbf{0}}_{\mathcal{G}/S})$, the functor $T \mapsto \mathrm{Hom}_{T\text{-rings}}(\mathcal{Y} \times_S T, E^{\mathbf{0}}_{\mathcal{G}/S} \times_S T)$ is representable by a scheme $H_{S\text{-rings}}(\mathcal{Y}, E^{\mathbf{0}}_{\mathcal{G}/S})$, where ring homomorphisms are supposed to take $1_{\mathcal{Y}}$ to the identity map $\mathrm{id}_{\mathcal{G}}$ of $\mathcal{G}$.

Recall the ring $O = O_D$ in the corollary. Let $\underline{O}_{/S}$ be the constant ring scheme with fiber O; so, $\underline{O}(T) = O^{\pi_0(T)}$ for the set $\pi_0(T)$ of connected components of T. Since each connected component of $\underline{O}$ is isomorphic to S, we can apply the above argument, getting an S-scheme $H_{S\text{-rings}}(\underline{O}, E^{\mathbf{0}}_{\mathcal{G}/S})$. Since the endomorphism ring of an abelian scheme over a base is torsion-free (see Theorem 8.7), any nontrivial $\phi \in \mathrm{Hom}_{T\text{-rings}}(\underline{O}_T, E^{\mathbf{0}}_{\mathcal{G}/S} \times_S T)$ for $\underline{O}_T = \underline{O} \times_S T$ is an injection, because O_D modulo a nonzero two-sided ideal is a finite ring (simplicity of D). Thus we have an S-scheme $E^{O}_{\mathcal{G}/S}$ representing the functor: $T \mapsto \{\theta : O \hookrightarrow \mathrm{End}(\mathcal{G}_T/T)\}$, where θ is a ring embedding.

We verify projectivity of each connected component E of $E^{O}_{\mathcal{G}/S}$ by the valuative criterion. Let V be a valuation ring with fraction field K. Suppose a K-point f of E is over a V-point $f \in S$. As we already remarked, $\theta_f : O_D \hookrightarrow \mathrm{End}(\mathcal{G}_K)$ extends to an algebra homomorphism $O_D \hookrightarrow \mathrm{End}(\mathcal{G}_V)$. Then f extends to a V-point of E. By the valuative criterion, E is proper over S. We already know that E is quasi-projective and hence is projective over S.

We now verify finiteness of E over S. Take a geometric point $s \in S$. Since O_D is generated over $\mathbb{Z}$ (as a ring) by finitely many elements $x_1, \ldots, x_m$ invertible in D, $\theta_s : O \hookrightarrow \mathrm{End}(\mathcal{G}_{s/s})$ is determined by the values at the generators $x_1, \ldots, x_m$. Suppose that θ_s gives rise to a geometric point of E. Write $[\theta(x_j)]$ (resp. $[x_j]$) for the connected component of $E^{\mathbf{0}}_{\mathcal{G}/S}$ (resp. $\underline{O}$) containing $\theta_s(x_j)$ (resp. x_j). We have a morphism $E \to [x_j] \times_S [\theta(x_j)]$ by sending $\theta \in E(S')$ to the S'-point $(x_j, \theta(x_j))$ of the fibered product. Since θ is determined by the values at x_j, this functorial map induces an immersion of E into $([x_1] \times_S [\theta(x_1)]) \times_S \cdots \times_S ([x_m] \times_S [\theta(x_m)])$. Since E is projective, this immersion is a closed immersion. Since $\underline{O}$ is constant, the above fibered product is isomorphic to $[\theta(x_1)] \times_S \cdots \times_S [\theta(x_m)]$. Since $\theta(x_j)$ is an isogeny $(x_j \in D^\times)$, $[\theta(x_j)]$ is finite over S. This shows the finiteness of $E_{/S}$. Since we may assume that $\theta(x_j)$ is an isogeny of degree prime to p for a given prime

p for all j, $[\theta(x_j)]$ is unramified over its image in S if S is a scheme over a discrete valuation ring of residual characteristic $p > 0$. This shows that E is unramified over its image in S. □

We can extend the above result to a semi-simple D by an induction of the number of simple components of D, which is slightly more technical.

We fix a base Dedekind domain $\mathcal{V}$ in a number field. Suppose that for each maximal ideal P of $\mathcal{V}$ and its residue field $k(P)$, $O_D \otimes_{\mathcal{V}} k(P)$ is semi-simple. We take an $O_D \otimes_{\mathbb{Z}} \mathcal{V}$-module V which is $\mathcal{V}$-locally free of finite rank n. Choose a base $v_1, \ldots, v_n$ of $V \otimes_{\mathcal{V}} k(0)$ over the field of fractions $k(0)$ of $\mathcal{V}$, and write $\rho(\alpha) \in M_n(\mathcal{V})$ for the matrix representation of $\alpha \in O_D$ given by $(\alpha v_1, \ldots, \alpha v_n) = (v_1, \ldots, v_n)\rho(\alpha)$. For each $\alpha \in O_D$, we define the characteristic polynomial $P(\alpha; t) = \det(t \cdot 1_n - \rho(\alpha)) \in \mathcal{V}[t]$. Suppose that for a field which is a $\mathcal{V}$-algebra, we have an $O_D \otimes k$-module V_1 of dimension n on which $\alpha \in O_D$ has characteristic polynomial $P(\alpha; t)$ for all $\alpha \in O_D$. Since representations of a semi-simple algebra over a field are characterized by the trace of representations (see [MFG] Theorem 2.6 or [BAL] VIII.6.5, VIII.13.3), we find that $V_1 \cong V \otimes_{\mathcal{V}} k$ as $O_D \otimes_{\mathbb{Z}} \mathcal{V}$-module.

For each semi-abelian scheme $\mathcal{G}_{/S}$ with $O_D \hookrightarrow \mathrm{End}(\mathcal{G}_{/S})$ as above, the tangent space at the origin $Lie(\mathcal{G})$ is a module over $O_D \otimes_{\mathbb{Z}} \mathcal{O}_S$. Since $\mathcal{G}$ is smooth over S. $Lie(\mathcal{G})$ is a locally free $\mathcal{O}_S$-module of rank $\dim_S \mathcal{G}$ (cf. [NMD] Proposition 2.2.5). If S is $Spec(k)$ for a field k and if $O_D \otimes_{\mathbb{Z}} k$ is semi-simple, there are only finitely many isomorphism classes of $O_D \otimes_{\mathbb{Z}} k$-modules of a given dimension (cf. [BAL] VIII.6.5).

Let S be a $\mathcal{V}$-scheme and $\mathcal{G}_{/S}$ be a (generically abelian) semi-abelian scheme. We consider for an S-scheme T the following condition.

(det) We have $\theta : O_D \hookrightarrow \mathrm{End}(\mathcal{G}_{T/T})$, and the characteristic polynomial of $\theta(\alpha)$ for all $\alpha \in O_D$ on $Lie(\mathcal{G}_T)$ over $\mathcal{O}_T$ is given by the image of $P(\alpha; t)$ in $\mathcal{O}_T[t]$.

We consider the functor from S-schemes into $SETS$ given by

$$T \mapsto \left[(\mathcal{G}_T, \theta : O_D \hookrightarrow \mathrm{End}(\mathcal{G}_T)) \;\middle|\; \theta \text{ satisfies (det)}\right]. \tag{6.6}$$

Since the property (det) is compatible with base change, the above functor is a well-defined subfunctor of the functor in Corollary 6.10.

Supposing that $\mathcal{G}_{/S}$ is an abelian scheme with a symmetric polarization $\lambda : \mathcal{G} \to {}^t\mathcal{G}$ (${}^t\lambda = \lambda$), we introduce a subfunctor of (6.6): We have the Rosati involution $\alpha \mapsto \alpha^* := \lambda^{-1} \circ {}^t\alpha \circ \lambda$ of $\mathrm{End}^{\mathbb{Q}}(\mathcal{G}_{/S}) = \mathrm{End}(\mathcal{G}_{/S}) \otimes_{\mathbb{Z}} \mathbb{Q}$. We suppose having an involution ρ of D and consider the following subfunctor

$$T \mapsto \left[(\mathcal{G}_T, \lambda_T, \theta) \text{ as in (6.6)} \middle| \theta(\alpha^\rho) = \theta(\alpha)^* \text{ for all } \alpha \in O_D\right]. \tag{6.7}$$

On each irreducible component I of the scheme representing the functor of Corollary 6.10, the characteristic polynomial of $\alpha \in O_D$ on $Lie(\mathcal{G}_\eta)$ for the generic fiber $\mathcal{G}_\eta$ gives a polynomial $P_I(\alpha; t) \in \mathcal{O}_I[t]$. For all points $x \in I$,

the specialization of $P_I(\alpha;t)$ at x gives the characteristic polynomial of α on $Lie(\mathcal{G}_x)$, because $Lie(\mathcal{G}_{/I})$ is locally $\mathcal{O}_I$-free of rank $\dim_S \mathcal{G}$. The functor in (6.6) is therefore represented by the union of irreducible components (characterized by (det)) of the scheme representing the functor in Corollary 6.10. In particular, the functor in (6.6) is representable by a scheme M over S whose connected components are finite over S.

We now study the representability of the functor in (6.7). We have an involution $[\rho]$ of the functor of (6.6) induced by $(\mathcal{G}_T, \theta) \mapsto (\mathcal{G}_T, * \circ \theta \circ \rho)$. Thus $[\rho]$ acts on the S-scheme M representing (6.6) as an involution (still denoted by $[\rho]$). The fixed points by this involution $[\rho]$ form a closed subscheme $M_\rho = M \times_{M, 1_M, [\rho]} M$ of M. Over M_ρ, θ satisfies the requirement $\theta(\alpha^\rho) = \theta(\alpha)^*$; so, M_ρ represents the functor (6.7) (although M_ρ could be empty depending on the choice of ρ). Thus we get

Corollary 6.11 *Let the notation and the assumption be as in* Corollary 6.10. *Suppose that S is a scheme over a Dedekind domain $\mathcal{V}$ in a number field. Suppose that $O_D \otimes_{\mathcal{V}} \mathcal{V}/\mathfrak{m}$ is semi-simple for every maximal ideal $\mathfrak{m}$ of $\mathcal{V}$. Then we have*

(1) *The functor of* (6.6) *is represented by a scheme over S;*
(2) *If $\mathcal{G}_{/S}$ is an abelian scheme, the functor* (6.7) *for an involution ρ of D is represented by a scheme over S.*

Each connected component of the scheme representing (6.6) *and* (6.7) *is finite over S, and if $\mathcal{V}$ is a valuation ring, it is finite unramified over S.*

6.2 Quotients by $PGL(n)$

A given abelian scheme A with a polarization is projective by the existence of an ample line bundle $\mathcal{L}_{/A}$ giving the polarization. Fixing a type of polarization, we may bound the rank of $H^0(A, \mathcal{L})$. Thus we get an embedding $A \hookrightarrow \mathbf{P}^m$ for a fixed sufficiently large m independent of A (once the dimension of A and the degree of the polarization are fixed). Hence abelian schemes over S with polarization give a subset $H(S)$ of $\mathrm{Hilb}_{\mathbf{P}^m}(S)$. By changing the base of $H^0(A, \mathcal{L})$, $PGL(m+1)$ acts on this subset preserving the isomorphism classes of abelian schemes. Thus to construct the moduli scheme of abelian varieties, we need to make the quotient space by $PGL(m+1)$. Although this is an oversimplified description of the construction of the moduli, it is obvious from this that we need to study the quotient of quasi-projective schemes by projective linear transformations. Here we reproduce an ingenious way of Mumford and Tate (in [GIT] Chapters 3–4), constructing a large open set U_{st} in $(\mathbf{P}^m)^M$, which has a structure of a $PGL(m+1)$-torsor under the diagonal action of $PGL(m+1)$ (see [GME] 1.8.3 for torsors). We later show that the subset H is actually a scheme and can be embedded into $(\mathbf{P}^m)^M$ for sufficiently large m and M preserving the action of $PGL(m+1)$; so, the base of the torsor H gives the desired moduli space.

6.2.1 Line Bundles on Projective Spaces

We want to determine line bundles on a given projective space $\mathbf{P}^n$ up to isomorphisms. For any given scheme S, we write $\mathrm{Pic}(S)$ for the set of isomorphism classes of invertible sheaves on S. For any given morphism $f : X \to S$, we consider the relative Picard functor $\mathrm{Pic}_{X/S}(T) = \mathrm{Pic}(X_T)/f_T^*(\mathrm{Pic}(T))$, where $f_T : X_T = X \times_S T \to T$ is the base change of f. Let us write $\mathbf{P}^n_{/\mathbb{Z}}$ for the projective space $Proj(\mathbb{Z}[X_0, \ldots, X_n])$. Since $\mathbf{P}^n$ has a section $\mathbf{0} = (1, 0, \ldots, 0)$, we can identify $\mathrm{Pic}_{\mathbf{P}^n/\mathbb{Z}}$ with $\mathrm{Ker}(\mathbf{0}^*)$; so, $\mathrm{Pic}_{\mathbf{P}^n/\mathbb{Z}}$ is local. In other words, $\mathcal{L} \in \mathrm{Pic}_{\mathbf{P}^n/\mathbb{Z}}(T)$ is determined by its restriction $\mathcal{L}|_{T_\iota}$ for any faithfully flat covering $\{T_\iota \to T\}_\iota$ of finite presentation. Suppose that $T = Spec(R)$. We take a graded $R[X_0, \ldots, X_n]$-module M with the associated sheaf $\widetilde{M} \in \mathrm{Pic}(\mathbf{P}^n_T)$. Then $M = \bigoplus_{n \geq -m} M_n$ and hence fiber-by-fiber, H^1 of $\widetilde{M} \otimes \mathcal{O}(m)$ vanishes (see [ALG] III.5). For $\pi : \mathbf{P}^n_T \to T$, $\pi_*(\widetilde{M} \otimes \mathcal{O}(m))$ is a locally free sheaf ([ALG] Theorem III.12.11). Since $\widetilde{M}_t$ for a closed point $t \in T$ is isomorphic to $\mathcal{O}(j)$ for some j, we may assume that $j = 0$ (suitably choosing m) and hence $\pi_*(\widetilde{M} \otimes \mathcal{O}(m)) \cong \mathcal{L}$ for an invertible sheaf $\mathcal{L}_{/T}$. Then $\pi_*(\widetilde{M} \otimes \mathcal{O}(m) \otimes \pi^*\mathcal{L}^{-1})$ is canonically isomorphic to $\mathcal{O}_T$, and hence, fiber-by-fiber trivial, which implies that it is globally trivial. This shows that $\widetilde{M} \cong \pi^*(\mathcal{L}) \otimes \mathcal{O}(m)$ for a suitable $\mathcal{L} \in \mathrm{Pic}(T)$ and $m \in \mathbb{Z}$, and

$$\mathrm{Pic}_{\mathbf{P}^n/\mathbb{Z}}(T) \cong \mathbb{Z}^{\pi_0(T)}. \tag{6.8}$$

6.2.2 Automorphism Group of a Projective Space

We consider the affine scheme $\mathbf{A}^n_{/\mathbb{Z}} = Spec(\mathbb{Z}[X_1, \ldots, X_n])$ on which the algebraic group $GL(n)$ acts by linear transformation. Recall that $GL(n)_{/\mathbb{Z}}$ is the open subscheme $Spec(\mathbb{Z}[X_{\imath\jmath}, \frac{1}{\det(X_{\imath\jmath})}])$ of $\mathbf{A}^{n^2}_{/\mathbb{Z}} = Spec(\mathbb{Z}[X_{\imath\jmath}])$. Similarly, we define $PGL(n)$ by the open subscheme of $\mathbf{P}^{n^2-1}$ given by removing the closed subscheme defined by $Proj(\mathbb{Z}[X_{\imath\jmath}]/(\det(X_{\imath\jmath})))$. Then $GL(n)$ and $PGL(n)$ are group schemes, and the sequence: $1 \to \mathbb{G}_m \to GL(n) \to PGL(n) \to 1$ is exact as *fppf* group sheaves (see Section 8.2.1 for a definition of *fppf* sheaves). The action induces automorphisms of $\mathbf{P}^{n-1}$.

Proposition 6.12 *If $S = Spec(R)$ is connected, $PGL_n(R) = \mathrm{Aut}(\mathbf{P}^{n-1}_{/R})$.*

Proof. By definition, we have $PGL_n(R) \subset \mathrm{Aut}(\mathbf{P}^{n-1}_{/R})$; so, we need to prove that $\alpha \in \mathrm{Aut}(\mathbf{P}^{n-1}_{/R})$ is induced by an element of $PGL(n)$. By (6.8), we see that $\alpha^*(\mathcal{O}(1)) = \mathcal{O}(m) \otimes \pi^*(\mathcal{L})$. Since $\mathcal{O}(1)$ among invertible sheaves $\mathcal{O}(m) \otimes \pi^*(\mathcal{L})$ is characterized by the fact that $H^q(\mathbf{P}^{n-1}_{/R}, \mathcal{O}(1))$ vanishes if $q \neq 0$ and

$$H^0(\mathbf{P}^{n-1}_{/R}, \mathcal{O}(1)) \cong R^n,$$

the same fact for $\alpha^*(\mathcal{O}(1))$ shows that $\mathcal{L}$ is trivial and $m = 1$. Thus α induces an automorphism of $\mathcal{O}(1)$, which is an element of $GL_n(R)$. □

6.2.3 Quotient of a Product of Projective Spaces

Following an ingenious trick of Mumford and Tate [GIT] Chapter 3, we now create a large open subset U_{st} of $\mathbf{P} = (\mathbf{P}^n)^{m+1}$ (for $m = n + \nu$ with $\nu > 0$) stable under the diagonal action of $PGL(n+1)$ which has a geometric quotient Z. We write $X_j^{(i)}$ for the homogeneous coordinates of the ith factor $\mathbf{P}^n$ in $(\mathbf{P}^n)^{m+1}$; thus, $i = 0, 1, \ldots, m$ and $j = 0, 1, \ldots, n$. We define $\mathcal{L}_i$ by the pullback of $\mathcal{O}_{\mathbf{P}^n}(1)$ to the ith factor of $\mathbf{P} = (\mathbf{P}^n)^{m+1}$.

Definition 6.13 For a subsequence $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_n)$ of $(0, 1, 2, \ldots, m)$, we define $D_\alpha = \det\left(X_i^{(\alpha_j)}\right)_{0 \le i,j \le n} \in \Gamma(\mathbf{P}, \mathcal{L}_\alpha)$, where $\mathcal{L}_\alpha = \bigotimes_{j=0}^n \mathcal{L}_{\alpha_j}$.

For each α_j, $GL(n+1)$ acts on the column vector ${}^t(X_0^{(\alpha_j)}, X_1^{(\alpha_j)}, \ldots, X_n^{(\alpha_j)})$ by matrix multiplication; thus, $D_\alpha(\gamma \circ x) = \det(\gamma) D_\alpha(x)$ for $\gamma \in GL(n+1)$. Thus D_α is invariant under $GL(n+1)$ up to units.

Proposition 6.14 *Let $\{S_1, S_2, \ldots, S_\nu\}$ be a collection of subsets of the set $\{0, 1, \ldots, n\}$ such that*

(1) *S_j and $\bigcup_{i=1}^{j-1} S_i$ have only one common number $\mu(j)$ for all $2 \le j \le \nu$, and we define $\mu(1)$ to be the least integer in S_1;*
(2) *$\{0, 1, \ldots, n\} = \bigcup_{j=1}^{\nu} S_j$.*

Define an affine open subscheme $U \subset \mathbf{P}$ by

(a) *$D_{\mathbf{1}} \neq 0$ for $\mathbf{1} = (0, 1, \ldots, n)$;*
(b) *$D_{\mathbf{1}(i)} \neq 0$ for all $i \in S_k$ and for all $k = 1, 2, \ldots, \nu$, where*

$$\mathbf{1}(i) = (\mathbf{1} - \{i\}) \cup \{n + k\}.$$

Then we have a $PGL(n+1)$-equivariant isomorphism:

$$U \cong PGL(n+1) \times_{Spec(\mathbb{Z})} \mathbf{A}^{n\nu - n}$$

for the affine space $\mathbf{A}^{n\nu-n}$ of dimension $n\nu - n$, where $PGL(n+1)$ acts on the right-hand side by multiplication on the left factor $PGL(n+1)$.

We reproduce the proof given in [GIT] Chapter 3, Section 1.

Proof. Since at least one of the equations $D_{\mathbf{1}(i)}$ involves variables of a given piece of $\mathbf{P}^n$, the open subscheme U is affine, because one could think of $\mathbf{P}$ as the quotient of $(\mathbf{A}^{n+1} - \{\mathbf{0}\})^{m+1}$ by $\mathbf{G}_m^{m+1}$ by component-wise multiplication on each piece $\mathbf{A}^{n+1} - \{\mathbf{0}\}$ (a sort of "*Proj*" of the graded algebra indexed by multiple degree with values in $\mathbb{Z}^{m+1}$).

Identify $\mathbf{A}^{n\nu-n}$ with $Spec(R)$ for

$$R = \frac{\mathbb{Z}[x_0^{(n+1)}, \ldots, x_n^{(n+1)}; x_0^{(n+2)}, \ldots, x_n^{(n+2)}; \ldots; x_0^{(n+\nu)}, \ldots, x_n^{(n+\nu)}]}{(x_i^{(n+k)} - 1 | i \in S_k,\ k = 1, 2, \ldots, \nu)}.$$

To make this legitimate, we count the number of relations: $x_i^{(n+k)} - 1$ for each $i \in \bigcup_j S_j = \{0, 1, 2, \dots, n\}$. If $i = \mu(k)$, the index i produces two equations for $k = 2, \dots, \nu$, and the number of equations is $(n+1) + (\nu - 1) = n + \nu$. The above spectrum is a linear subspace in $(n+1)\nu$-dimensional affine space defined by $n + \nu$ independent linear equations; so, it is isomorphic to $\mathbf{A}^{n\nu - n}$.

We then define an embedding $\iota : \mathbf{A}^{n\nu-n} \hookrightarrow U$ as follows. For a graded algebra $A = \bigoplus_k A_k$ and a homogeneous element $a \in A_+ = \oplus_{k>0} A_k$, we write $D_+(a) = Spec(A[\frac{1}{a}]_0) \subset Proj(A)$, where the subscript 0 indicates the degree 0 component. Writing $p_i : \mathbf{P} \to \mathbf{P}^n$ for the ith projection, the composite $\iota_i = p_i \circ \iota$ which has values in $D_+(Y)$ for $Y \in \{X_0^{(i)}, \dots, X_n^{(i)}\}$ is determined by an algebra homomorphism $\iota_i^* : \mathbb{Z}[X_0^{(i)}/Y, \dots, X_n^{(i)}/Y] \to R$.

When $i \le n$, we take Y given by $Y = X_i^{(i)}$, and ι_i^* is constant taking $X_j^{(i)}/X_i^{(i)}$ to the Kronecker symbol δ_{ij}. When $i > n$, then we have $i - n = k$ for some $k = 1, 2, \dots, \nu$, and we choose $i_i : Spec(R) \to \mathbf{P}^n$ sending the coordinate $(x_0^{(n+1)}, \dots, x_n^{(n+1)}; x_0^{(n+2)}, \dots, x_n^{(n+2)}; \dots; x_0^{(n+\nu)}, \dots, x_n^{(n+\nu)})$ of $Spec(R)$ to the homogeneous coordinate $(x_0^{(i)}, \dots, x_n^{(i)})$ of $\mathbf{P}^n$. This is legitimate because for $j \in S_k$, we have $x_j^{(i)} = 1$ in $Spec(R)$. This defines a closed immersion of $Spec(R)$ into U, and $\iota_i^*(\mathcal{O}_{\mathbf{P}^n}(1))$ is the structure sheaf $\widetilde{R}$ of $Spec(R)$.

Since D_α is invariant (up to the scalar) under the action of $GL(n+1)$, U is stable under $PGL(n+1)$, and we can extend ι to $PGL(n+1) \times Spec(R) \xrightarrow{\iota} U$ by $\iota(g, x) = g(\iota(x))$.

To define the inverse ϕ of ι, for each given $x \in U$, we need to show that we can find a unique $g \in PGL(n+1)$ so that for $i \le n$, $p_i(g(x))$ is the origin $\mathbf{0}$ in $D_+(X_i^{(i)}) = \mathbf{A}^n$ with homogeneous coordinate $(0, \dots, 0, \overset{i}{1}, 0, \dots, 0)$. The uniqueness, up to right multiplication by diagonal matrices, is plain, because the homogeneous coordinates of $p_i(x)$ for $0 \le i \le n$ forms an $(n+1) \times (n+1)$ matrix g with determinant $D_{\mathbf{1}}(x)$ which is invertible on U.

We need to check that $g^{-1}x$ is in the image of ι and to be able to choose g without ambiguity of diagonal matrices. For that purpose, we introduce some notation. We have defined $S_k \cap (S_{k-1} \cup \cdots \cup S_1) = \{\mu(k)\}$ for $k \ge 2$, and $\mu(1)$ is the least integer in S_1. We write $\kappa(i)$ for each $i = 0, i, \dots, n$ the least integer with $i \in S_{\kappa(i)}$. By definition, i and $\mu(\kappa(i))$ are both in $S_{\kappa(i)}$. Repeating this, for $\ell = \mu \circ \kappa$, $\ell^j(i) \in S_{\kappa(\ell^{j-1}(i))}$. Writing $i \succ j$ if $\kappa(i) > \kappa(j)$, we have

$$i \succ \ell(i) \succ \ell^2(i) \succ \ell^3(i) \succ \cdots \succ \ell^k(i) = \mu(1) \quad \text{for some } k.$$

Then we define $(\lambda_0(x), \lambda_1(x), \dots, \lambda_n(x))$, by induction on j,

$$\lambda_{\mu(1)}(x) = 1 \text{ if } j = 1 \text{ and } \lambda_j(x) = \lambda_{\ell(j)}(x) \frac{D_{\mathbf{1}(\ell(j))}(x)}{D_{\mathbf{1}(j)}(x)} \text{ if } j > 1.$$

Over U, $\lambda_j(x)$ is a well-defined section of $\mathcal{L}_j \otimes \mathcal{L}_{\mu(1)}^{-1}$.

We finally define $\phi : U \to PGL(n+1) \times Spec(R)$ as follows. Writing two projections of ϕ as $\phi_P : U \to PGL(n+1) \subset Proj(\mathbb{Z}[T_{ij}]_{0 \le i,j \le n})$ and

$\phi_R : U \to Spec(R)$, we define ϕ_P by $\phi_P^*(\mathcal{O}_{PGL}(1)) = \mathcal{L}_{\mu(1)}$ and $\phi_P^*(T_{ij}) = X_i^{(j)} \otimes \lambda_j^{-1}$; here $\mathcal{O}_{PGL}(1)$ is the restriction of $\mathcal{O}(1)$ of $Proj(\mathbb{Z}[T_{ij}]_{0\le i,j\le n})$ to the affine open subset $PGL(n+1)$. By definition, $\phi(x) = g \cdot \mathrm{diag}[\lambda_0, \ldots, \lambda_n]^{-1}$ for $g = (p_0(x), \ldots, p_n(x))$; so, we have now specified g without any ambiguity. The point here is that we have now made the choice unique (not just up to multiple by diagonal matrices). We then define

$$\phi_R^*(x_i^{(n+k)})(x) = \frac{D_{\mathbf{1}(i)}(x)}{D_{\mathbf{1}(\mu(k))}(x)} \cdot \frac{\lambda_i(x)}{\lambda_{\mu(k)}(x)}.$$

It is easy to check that ϕ_R has values in $Spec(R)$. Since $\phi_P(x)$ and $\phi_R(x)$ contain all the information necessary to compute the coordinates of x, it is an immersion. By a simple combinatorial computation, basically by definition, we check that $\phi \circ \iota$ is the identity map of $PGL(n+1) \times Spec(R)$; so, we conclude that ι is a surjective isomorphism. □

Definition 6.15 $U_{st} \subset \mathbf{P}$ is an open subscheme whose geometric points $x = (x^{(0)}, \ldots, x^{(m)})$ $(m = n + \nu)$ with $x^{(j)} \in \mathbf{P}^n$ satisfy

$$\frac{\text{number of points } x^{(i)} \text{ in } L}{m+1} < \frac{\dim L + 1}{n+1}$$

for every proper linear subspace $L \subset \mathbf{P}^n(k(x))$.

A point $x \in \mathbf{P}$ is called *stable* if $x \in U_{st}$.

We need to verify that U_{st} is an open subscheme. Let $I \subset \{0, 1, \ldots, m\}$ and choose an integer n_0 with $0 \le n_0 \le n-1$. Suppose that

$$\frac{|I|}{m+1} \ge \frac{n_0+1}{n+1}.$$

Then we see $\mathbf{P} - U_{st} = \bigcup_{I,n_0} V(I, n_0)$, where

$$V(I, n_0) = \left\{ x \in \mathbf{P} \middle| \left(x^{(i)}\right)_{i\in I} \text{ spans a linear subspace of dimension} \le n_0 \right\}.$$

Since the matrix X made of homogeneous coordinates of $x^{(i)}$ for $x \in V(I, n_0)$ has rank at most n_0, all the $(n_0+2) \times (n_0+2)$-minors of X have determinant 0. Thus $V(I, n_0)$ is a closed subscheme and, hence, U_{st} is an open subscheme.

Proposition 6.16 *A geometric point $x \in U_{st}$ if and only if there are integers $N \ge N_0 > 0$ and monomials $P_0, \ldots, P_m$ of D_α as in* Definition 6.13 *such that, for $0 \le i \le m$, $P_i \in \Gamma(\mathbf{P}, \mathcal{L}_0^N \otimes \cdots \otimes \mathcal{L}_i^{N-N_0} \otimes \cdots \otimes \mathcal{L}_m^N)$ and $P_i(x) \ne 0$.*

We repeat the proof of Mumford ([GIT] Proposition 3.6).

Proof. We associate with the monomials $D = \prod_\alpha D_\alpha^{e(\alpha)}$ an integral point $d(D) \in \mathbb{R}^{m+1} = V$ in the following way. the ith coordinate of $d(D)$ is

$\sum_{\alpha \ni \imath} e(\alpha)$. For $x \in \mathbf{P}$, define $C_x \subset V$ to be the simplicial cone spanned by $d(D_\alpha)$ for $D_\alpha(x) \neq 0$; that is, C_x is made up of a positive linear combination of $d(D_\alpha)$ with $D_\alpha(x) \neq 0$. Thus $d(D) \in C_x \iff D(x) \neq 0$. Writing $\mathbf{1} = (1, 1, \ldots, 1) \in V$, we see that $\mathbf{1} \in C_x \iff$ we have $N\mathbf{1} - v \in C_x$ with a sufficiently large N for any given $v \in V$. Thus taking $v = N_0(0, \ldots, 0, 1, 0, \ldots, 0)$ for N_0 in the proposition, we need to show that $\mathbf{1} \in C_x \iff x \in U_{st}$. Then by the lemma following this proof, we need to prove $x \in U_{st}$ if and only if for any given partition $\{0, 1, \ldots, m\} = I \sqcup J$, there exists α such that $D_\alpha(x) \neq 0$ and

$$|I| \cdot |\alpha \cap J| < |J| \cdot |\alpha \cap I|. \tag{6.9}$$

Suppose $\{x^{(\imath)}\}_{\imath \in I}$ spans a linear subspace L of $\mathbf{P}^n$. Then $D_\alpha(x) \neq 0$ implies the maximum number of $\alpha_\imath$ in I has to be $\dim L$; so, (6.9) holds (in addition to $D_\alpha(x) \neq 0$) if and only if

$$|I| \cdot (n - \dim L) \leq |I| \cdot |J \cap \alpha| < |J| \cdot |\alpha \cap I| \leq (m + 1 - |I|)(\dim L + 1)$$

which is equivalent to the inequality in Definition 6.15 of stability of the point x. □

We need to prove the following lemma.

Lemma 6.17 *Let the notation be as in* Proposition 6.16 *and its proof. We have* $\mathbf{1} \notin C_x$ *if and only if there exists a partition of sets:* $\{0, 1, \ldots, m\} = I \sqcup J$ *such that the inequality* (6.9) *fails to hold for any* α *with* $D_\alpha(x) \neq 0$.

Proof. We have $\mathbf{1} \notin C_x$ if and only if there exists an integral linear form $\ell : \mathbb{Z}^{m+1} \to \mathbb{Z}$ such that

$$\ell(\mathbf{1}) = 0 \text{ and } \ell(d(D_\alpha)) \geq 0 \text{ for all } \alpha \text{ with } D_\alpha(x) \neq 0. \tag{6.10}$$

Since $\mathbf{1}$ is fixed by any permutation of coordinates, we may assume that $\lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_m$ if we write $\ell(v) = \sum_\imath \lambda_\imath v_\imath$ for $v = (v_\imath) \in \mathbb{R}^{m+1} = V$. Then the condition (6.10) is verified for ℓ given by

$$\lambda_0 = \lambda_1 = \cdots = \lambda_\beta = -(m - \beta) \text{ and } \lambda_{\beta+1} = \cdots = \lambda_m = \beta + 1.$$

Once this is proven, we just take $I = \{\beta + 1, \ldots, m\}$ and $J = \{0, 1, \ldots, \beta\}$, and the assertion follows.

We prove that the above expression of ℓ is possible if (6.10) holds. We may assume that $D_\alpha(x) \neq 0$ and $\ell(D_\alpha) = \sum_\imath \lambda_{\alpha_\imath}$ is minimal; so, we may assume that $x^{(0)}, \ldots, x^{(\beta_j - 1)}$ span a linear subspace of dimension $j - 1$ but $x^{(\beta_j)}$ does not lie in this linear subspace for all $j = 1, 2, \ldots, n$. Thus $x^{(0)}, \ldots, x^{(\beta_n)}$ span entire $\mathbf{P}^n$, and hence $D_\beta(x) \neq 0$ for $\beta = (0, \beta_1, \ldots, \beta_n)$, and

$$D_\alpha(x) \neq 0 \iff \alpha_0 \geq 0,\ \alpha_1 \geq \beta_1, \ldots, \alpha_n \geq \beta_n.$$

Hence the condition (6.10) is equivalent to $\ell(d(D_\beta)) \geq 0$. Since the set of ℓs,

$$\mathcal{C} = \left\{ (\lambda_0, \lambda_1, \ldots, \lambda_m) \in \mathbb{Z}^{m+1} \;\middle|\; \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_m \text{ and } \sum \lambda_i = 0 \right\},$$

is convex, if $\ell(d(D_\beta)) \geq 0$ holds in the interior of $\mathcal{C}$, it has to be true at least for one face, which is determined by the partition $I \sqcup J = \{0, 1, \ldots, m\}$ so that $\lambda_i = -(m - \beta)$ if $i \in J$ and $\lambda_j = \beta + 1$ if $j \in J$. □

Corollary 6.18 *Let the notation be as in* Propositions 6.14 *and* 6.16. *Suppose that $x \in \mathbf{P} = (\mathbf{P}^n)^{m+1}$ (for $m > n$) is stable. Then there exists a sequence of subsets $\{S_1, S_2, \ldots, S_\nu\}$ for $\nu = m - n$ satisfying the conditions in* Proposition 6.14 *such that $x \in U \cap U_{st}$, where U is the open subscheme defined by the sequence as in the proposition. Moreover, there is an open neighborhood U_0 of x in $U \cap U_{st}$ defined by $P \neq 0$ for a monomial P in the D_αs that is a section of $(\mathcal{L}_0 \otimes \mathcal{L}_1 \otimes \cdots \otimes \mathcal{L}_m)^M$ for some integer $M > 0$.*

Proof. Choose $N > N_0 > 0$ and choose $P_0, \ldots, P_m$ as in Proposition 6.16. Let $Q = \bigotimes_{D_\alpha(x) \neq 0} D_\alpha$. Then $P = Q^{N_0} \otimes P_0^{r_0} \otimes \cdots \otimes P_m^{r_m}$ for a suitable r_j is a section as in the proposition. Then $P \neq 0$ defines an open neighborhood U_0 by Proposition 6.16. We now construct a sequence S_j of subsets of $\{0, 1, \ldots, n\}$ so that $x \in U$. Since x is stable, there do not exist disjoint proper linear subspaces L' and L'' of $\mathbf{P}^n(k(x))$ such that $x^{(i)}$ is either in L' or L'' for all i. Thus all the points $x^{(i)}$ cannot lie in a hyperplane; so, we may assume that $x^{(0)}, \ldots, x^{(n)}$ cannot be in a hyperplane (by reordering coordinates). Furthermore, we may assume that $x_j^{(i)} = \delta_{ij}$ by moving around by the action of $PGL(n+1)$. Define sets S'_k $(k = 1, 2, \ldots, \nu)$ by the set of integers i such that $D_{0,1,\ldots,i-1,i+1,\ldots,n,n+k}(x) \neq 0$; that is, $x^{(n+k)}$ is not in the hyperplane. We prove that there is no partition $\{0, 1, \ldots, n\} = T' \sqcup T''$ such that either $S'_k \subset T'$ or $S'_k \subset T''$ for all k. If this is proven, by shrinking S'_k a little if necessary, we can achieve the condition of Proposition 6.14.

Suppose on the contrary that $S'_k \subset T'$ or $S'_k \subset T''$ for all k. Define linear subspaces L' (resp. L'') by $X_i = 0$ for all $i \in T'$ (resp. $i \in T''$). Then every point $x^{(k)}$ is either in L' or L'', contradicting the stability of x. □

Covering U_{st} by the open sets $\{U_i\}$ as in Corollary 6.18, we have $U_i = PGL(n+1) \times Z_i$. Thus on $U_i \cap U_j$, $Z_i \times_{U_i} (U_i \cap U_j)$ is canonically isomorphic to $Z_j \times_{U_j} (U_i \cap U_j)$; so, the schemes Z_j glue into a scheme Z with projection $\pi : U_{st} \to Z$. By our construction, over $U = \bigsqcup_i U_i$, $U_{st} \times_Z U = PGL(n+1) \times Z$. Thus U_{st} is a $PGL(n+1)$-torsor (see [GME] Section 1.8.3 for a definition of torsor), and hence (Z, π) is the geometric quotient of U_{st} (see [GIT] Section 0.6 or [GME] Section 1.8.3 for the definition of geometric quotients). The geometric quotient always gives a categorical quotient (see [GIT] Proposition 0.1 or [GME] Proposition 1.8.1).

Thus we get

Theorem 6.19 (Mumford) *The pair (Z, π) is the geometric and categorical quotient of U_{st} by $PGL(n+1)$ and is quasi-projective.*

6.3 Mumford Moduli

We describe the Mumford construction of the moduli over $\mathbb{Z}$ of abelian schemes of dimension n with a given polarization of degree d^2.

6.3.1 Dual Abelian Scheme and Polarization

Let $A_{/S}$ be an abelian scheme. We consider the following Picard functor

$$\mathrm{Pic}_{A/S}(T) = \mathrm{Pic}(A_T)/f_T^* \mathrm{Pic}(T) = \mathrm{Ker}(\mathbf{0}_T^*)$$

for $f : T \to S$. It is known that $\mathrm{Pic}_{A/S}$ is represented by a (locally Noetherian) reduced group scheme (Grothendieck [FGA] Exp. 232, 1961/62, [ABV] Section 13, or [NMD] Chapter 8). Let ${}^tA$ be the identity-connected component $\mathrm{Pic}^\circ_{A/S}$ of the group scheme representing $\mathrm{Pic}_{A/S}$. Then ${}^tA_{/S}$ is an abelian scheme.

Let $\mathcal{L}^{univ} \in {}^tA({}^tA) = \mathrm{Pic}^\circ_{A\times_S {}^tA/{}^tA} = \mathrm{Ker}(\mathbf{0}^*_{A\times_S {}^tA})$ corresponding to the identity. Then the sheaf $\mathcal{L}^{univ}$ has the following universal properties:

- $\mathbf{0}^*_{A\times_S {}^tA}\mathcal{L}^{univ} = \mathcal{O}_{{}^tA}$;
- Let $T \to S$ be an S-scheme. For any invertible sheaf L on A_T algebraically equivalent to $\mathcal{O}_{A_T}$, there exists a unique morphism $\iota_L : T \to {}^tA$ such that $(id_A \times \iota_L)^*\mathcal{L}^{univ} = L$.

Let L be an invertible sheaf on A. For $x \in A$, we define the translation $T_x(y) = x + y$, which is an automorphism of A. Then $(T_x^*L)\otimes L^{-1}$ is an element in ${}^tA$, and we obtain a morphism $\Lambda(L) : A \to {}^tA$. This S-homomorphism is an *isogeny* (i.e., surjective) if and only if L is *ample* (i.e., choosing a base of $H^0(A, L^{\otimes n})$ for sufficiently large n, one can embed A into $\mathbf{P}^m_{/S}$ over S; see [ABV] Section 6). The degree of the polarization is defined to be the square root of the degree of the homomorphism $\Lambda(L)$.

There is another construction of $\Lambda(L)$. Consider $a^*(L) \otimes p_1^*(L) \otimes p_2^*(L)^{-1}$ as an invertible sheaf on $A_A = A \times_S A$, where $a : A \times A \to A$ is the addition of the group scheme A. Then this invertible sheaf induces an A-valued point of $\mathrm{Pic}_{A/S}(A)$, which factors through ${}^tA$, because at the identity, this sheaf specializes to the trivial invertible sheaf at the origin $\mathbf{0}$ (so, the image is in the connected component of $\mathrm{Pic}_{A/S}$). We claim this A-valued point of ${}^tA$ is actually $\Lambda(L)$. Indeed, specializing this sheaf to $(x : S \to A) \in A(S)$, we get $T_x^*(L) \otimes L^{-1} \otimes x^*(L)$, which is equivalent in $\mathrm{Pic}(A)/f^* \mathrm{Pic}(S)$ to $T_x^*L \otimes L^{-1}$.

6.3.2 Moduli Problem

Fix positive integers n, d, and N, and define a functor $\mathcal{A}_{d,N} : SCH_{/\mathbb{Z}[\frac{1}{N}]} \to SETS$ by

$$\mathcal{A}_{d,N}(S) = \left[(A, \phi_N : (\mathbb{Z}/N\mathbb{Z})^{2n} \cong A[N], \lambda)_{/S}\right],$$

where

(A1) $A_{/S}$ is an abelian scheme with $\dim_S A = n$,
(A2) ϕ_N is an isomorphism of group schemes over S, and
(A3) λ is a polarization, locally $\lambda = \Lambda(\mathcal{L})$ on S and $\deg(\lambda) = d^2$.

It is known (see [GIT] Proposition 6.10) that if λ is locally of the form $\Lambda(L)$, then 2λ is globally $\Lambda(L^{\Delta}(\lambda))$ for the invertible sheaf $L^{\Delta}(\lambda)$ given by $(1_A \times \lambda)^*(\mathcal{L}^{univ})$. Here is a theorem of Mumford.

Theorem 6.20 (Mumford) *There exists a quasi-projective scheme $\mathfrak{M}_{d,N}$ over $\mathbb{Z}[\frac{1}{N}]$ such that*

(1) *For any geometric point $s = Spec(k)$ of $Spec(\mathbb{Z}[\frac{1}{N}])$, $\mathcal{A}_{d,N}(s) \cong \mathfrak{M}_{d,N}(s)$;*
(2) *If $N \geq 3$, we have a universal object $(\mathbf{X}, \boldsymbol{\phi}_N, \boldsymbol{\lambda})_{/\mathfrak{M}_{d,N}}$ such that for each triple $(A, \phi_N, \lambda) \in \mathcal{A}_{d,N}(S)$ there exists a unique morphism $\iota : S \to \mathfrak{M}_{d,N}$ with $(A, \phi_N, \lambda) \cong (\mathbf{X}, \boldsymbol{\phi}, \boldsymbol{\lambda}) \times_{\mathfrak{M}_{d,N}} S = \iota^*(\mathbf{X}, \boldsymbol{\phi}, \boldsymbol{\lambda})$, and the above association $(A, \phi_N, \lambda)_{/S} \mapsto \iota$ induces a functorial isomorphism $\mathcal{A}_{d,N}(S) \cong \underline{\mathfrak{M}}_{d,N}(S)$, where $\underline{\mathfrak{M}}_{d,N}(S) = \mathrm{Hom}_{\mathbb{Z}[\frac{1}{N}]}(S, \mathfrak{M}_{d,N})$;*
(3) *If $N \geq 3$, $\mathfrak{M}_{d,N}$ is smooth over $\mathbb{Z}[\frac{1}{Nd}]$.*

When d is prime to N, the pairing $A[N] \times {}^tA[N] \to \mu_N$ composed with the polarization gives a self-duality $e_{A,N} : A[N] \times A[N] \to \mu_N$. Recall $J = \left(\begin{smallmatrix} 0 & -1_n \\ 1_n & 0 \end{smallmatrix}\right)$, fix a primitive Nth root of unity ζ_N, and impose one more compatibility condition in addition to (A1–3) to define a subfunctor $\mathcal{A}^J_{1,N} \subset \mathcal{A}_{1,N}$.

(A4) We have $e_{A,N}(\phi_N(x), \phi_N(y)) = \zeta_N^{{}^tx \cdot Jy}$ for all $x, y \in (\mathbb{Z}/N\mathbb{Z})^{2n}$.

Since $e_{A,N}$ pulled back by ϕ_N can be written as $(x, y) \mapsto \zeta_N^{{}^tx \cdot J'y}$ for a unique nondegenerate alternating matrix J', $\mathfrak{M}_N = \mathfrak{M}_{1,N}$ is a disjoint union of $\mathfrak{M}^{J'}_{1,N}$ classifying triples (A, ϕ_N, λ) satisfying (A1–3) and the condition obtained from (A4) replacing J by J'. Thus fixing $e = (1, 0, \ldots, 0)$ and $e' = (0, \ldots, 0, 1)$ in $(\mathbb{Z}/N\mathbb{Z})^{2n}$, we have a well-defined morphism $\mathcal{A}^J_{1,N} \to \underline{\mu}_N$ sending $(A, \phi_N, \lambda)_{/B}$ to $e_{A,N}(\phi_N(e), \phi_N(e')) \in \mu_N(B)$, which in turn induces a morphism of schemes $\mathfrak{M}^J_{1,N} \to \mathbb{Z}[\zeta_N, \frac{1}{N}]$ for the primitive root of unity $\zeta_N = e_{\mathcal{X},N}(e, e')$. Here $\mathcal{X}_{/\mathfrak{M}^J_N}$ is the universal abelian scheme. This shows that the scheme $\mathfrak{M}^J_N$ is well defined over $\mathbb{Z}[\zeta_N, \frac{1}{N}]$. Looking at the fiber over $\mathbb{C}$ of $\mathfrak{M}^J_{N/\mathbb{Z}[\zeta_N, \frac{1}{N}]}$, by an argument similar to the Hilbert modular case (see Section 4.1.3 and the proof of Theorem 6.26), we find $\mathfrak{M}^J_N(\mathbb{C}) \cong \Gamma(N)\backslash\mathfrak{H}_n$ for the Siegel upper half-space $\mathfrak{H}_n = \{{}^tz = z \in M_n(\mathbb{C}) | \, \mathrm{Im}(z) > 0\}$ and

$$\Gamma(N) = \left\{\alpha \in Sp_{2n}(\mathbb{Z}) \middle| \alpha \equiv 1_{2n} \mod N\right\}.$$

By the existence of a smooth toroidal compactification ([DAV]; see also Section 6.3.6 in the text), $\mathfrak{M}^J_N$ is geometrically irreducible over $\mathbb{Z}[\zeta_N, \frac{1}{N}]$ fiber-by-fiber.

Remark 6.21 If d is prime to N, choosing well a nondegenerate alternating matrix J' in $M_{2n}(\mathbb{Z}/N\mathbb{Z})$, we could impose

(A4′) We have $e_{A,N}(\phi_N(x), \phi_N(y)) = \zeta_N^{{}^t x \cdot J' y}$ for all $x, y \in (\mathbb{Z}/N\mathbb{Z})^{2n}$.

We then have a smooth geometrically irreducible scheme $\mathfrak{M}_{d,N}^{J'}$ over $\mathbb{Z}[\zeta_N, \frac{1}{dN}]$ representing the functor classifying triples (A, λ, ϕ_N) satisfying (A1–3) and (A4′). We apply Corollary 6.11 to the universal abelian scheme $\mathcal{X}_{/\mathfrak{M}_{d,N}^{J'}}$ and $O_D = O_F$ to verify the representability of the functor $\mathcal{E}_{\Gamma(N)}$ of $(\Gamma(N))$ in Section 4.1.2 by a scheme $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ as follows. First of all, the condition (det) applied to the integer ring O of F is equivalent to (rm4) in the definition of the AVRMs. Since we can choose the polarization ideal $\mathfrak{c}$ in its strict ideal class, we may assume the degree of the polarization and $\mathfrak{c}$ itself is prime to the given integer $N \geq 3$. Choosing a $\mathbb{Z}$-base of $O \oplus \mathfrak{c}^*$ and identifying $L = O \oplus \mathfrak{c}^*$ with $\mathbb{Z}^{2n}$ $(n = [F : \mathbb{Q}])$, the polarization pairing $L \wedge L \cong \mathfrak{c}^* \xrightarrow{\mathrm{Tr}} \mathbb{Q}$ gives rise to a pairing $(x, y) \mapsto {}^t x \cdot J' y$ for a suitably chosen alternating matrix $J' \in M_{2n}(\mathbb{Z})$ as above. By Corollary 6.11, $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ is realized as a scheme finite over $\mathfrak{M}_{d,N}^{J'}$; so, it is affine over $\mathfrak{M}_{d,N}^{J'}$ and is quasi-projective. All other Hilbert–Blumenthal moduli schemes $\mathfrak{M}(\mathfrak{c}, \Gamma)$ there for different level structure Γ can be obtained from $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ by making a quotient by a finite constant group. So the representability claimed in Section 4.1.2 is now verified.

We now give a sketch of Mumford's proof of the above theorem. Let $A_{/\mathcal{B}}$ be an abelian scheme over a ring $\mathcal{B}$. The key idea is that for a given very ample invertible sheaf $L_{/A}$, the embedded image of A under L in $\mathbf{P}^m$ for $m + 1 = \mathrm{rank}_{\mathcal{B}} H^0(A, L)$ is determined by the choice of basis b of $H^0(A, L)$. In other words, the image is transported to the other by an element of $PGL_{m+1}(\mathcal{B})$. Since for an abelian scheme, by the generalized Riemann–Roch theorem (see [ABV] Section 16), we can compute the Hilbert polynomial P of L, the moduli functor of (A, b) is a subfunctor of $\underline{\mathrm{Hilb}}^P_{\mathbf{P}^m}$. Proving that the image is a quasi-projective subscheme H of $\mathrm{Hilb}^P_{\mathbf{P}^m}$, the (coarse) modulus is constructed as $\mathfrak{M}_1 = PGL(m+1)\backslash H$. We give more details of this argument in the following subsections. The proof finishes at the end of Section 6.3.5.

6.3.3 Abelian Scheme with Linear Rigidification

Let $(A, \phi_N, \lambda)_{/S} \in \mathcal{A}_{d,N}(S)$. We consider the invertible sheaf $L_{/S}$ given by the direct image of $L^{\Delta}(\lambda)^3$, which is very ample, because $\mathcal{L}^3$ is very ample if $\mathcal{L}$ is ample (see [ABV] Section 17) and ampleness of $L^{\Delta}(\lambda)$ follows from $\Lambda(L^{\Delta}(\lambda)) = 2\lambda$ as we already remarked. Let $\mathbf{S}(L)$ be the symmetric algebra of L (as in Section 6.1.1) and write $\mathbf{P}(L)$ for the projective bundle $Proj(\mathbf{S}(L))$ which is a projective scheme over S locally isomorphic to $\mathbf{P}^m_{/S}$. A *linear rigidification* is an isomorphism $\iota : \mathbf{P}(L) \cong \mathbf{P}^m_{/S}$. Thus ι is determined by the choice of a base of L up to scalar multiplication. Since the very ample sheaf $L^{\Delta}(\lambda)^3$ on an abelian scheme satisfies ([ABV] Section 16)

- $L = f_*(L^{\Delta}(\lambda)^3)$ is locally free of finite rank $6^n d$, and
- $R^i f_*(L^{\Delta}(\lambda)^3) = 0$ if $i > 0$,

the formation of $f_*(L^{\Delta}(\lambda)^3)$ as above commutes with base change (cf. [ALG] III.12.11). Thus the association $S \mapsto \mathcal{A}'_{d,N}(S) = \left[(A, \phi_N, \lambda, \iota)_{/S}\right]$ is a well-defined contravariant functor. The embedding $I : A_{/S} \hookrightarrow \mathbf{P}(L) \xrightarrow{\iota} \mathbf{P}^m_{/S}$ determines the sheaf $L^{\Delta}(\lambda)^3 = I^*\mathcal{O}_{\mathbf{P}(L)}(1)$, which in turn determines λ because $\Lambda(L) = 6\lambda$ ($\mathrm{Pic}_{A/S}/{}^tA$ is torsion-free; see [ABV] Section 19 Corollary 2). Having ϕ_N is equivalent to having $2n$ (linearly independent) sections $\sigma_j = \phi_N(e_j)$ of A over S for the standard base $\{e_1, \ldots, e_{2n}\}$ of $(\mathbb{Z}/N\mathbb{Z})^{2n}$. We write $\sigma_0 = \mathbf{0}$ for the identity section $\mathbf{0}$ of A. We record here what we have seen.

Proposition 6.22 *The data $(A, \phi_N, \lambda, \iota)$ are completely determined by the embedding data $(I : A_{/S} \hookrightarrow \mathbf{P}^m_{/S}, \sigma_0, \sigma_1, \ldots, \sigma_{2n})$.*

In other words, defining another functor $\mathcal{H}_{d,N}(S) : \mathbb{Z}[\frac{1}{N}]\text{-}SCH \to SETS$ by

$$\mathcal{H}_{d,N}(S) = \left[(I : A_{/S} \hookrightarrow \mathbf{P}^m_{/S}, \sigma_0, \sigma_1, \ldots, \sigma_{2n})_{/S}\right],$$

we have an isomorphism of functors: $\mathcal{A}'_{d,N} \cong \mathcal{H}_{d,N}$. Here $\sigma_1, \ldots, \sigma_{2n}$ are supposed to give $2n$ linearly independent elements of $A[N](S)$.

6.3.4 Embedding into the Hilbert Scheme

For simplicity, we just write $\mathbf{P}$ for $\mathbf{P}^m_{\mathbb{Z}[\frac{1}{N}]}$. We write S_0 for $Spec(\mathbb{Z}[\frac{1}{N}])$. We consider the functor $\underline{\mathrm{Hilb}}^P_{\mathbf{P}}$ associating with each S the set of closed subschemes of $\mathbf{P}_{/S}$ flat over S with Hilbert polynomial P. As we have already seen, this functor is represented by a scheme $H = \mathrm{Hilb}^P_{\mathbf{P}}$. Write $Z \to H$ for the universal flat family inside $\mathbf{P}_{/H}$ with Hilbert polynomial P. For each subscheme $V \subset \mathbf{P}_{/S}$ flat over S having Hilbert polynomial P, we have a unique morphism $h : S \to H$ such that V is given by $S \times_H Z \subset \mathbf{P}_{/S}$ over S.

By the generalized Riemann–Roch theorem ([ABV] Section 16), the Hilbert polynomial of (A, L) (or of $I(A)$) is given by $P(T) = 6^n d \cdot T^n$. Thus the image $I(A)$ induces a unique S_0-morphism $h : S \to H$ such that $I(A) = S \times_H Z$ as subschemes of $\mathbf{P}$.

Let $H^k = \mathrm{Hilb}^{P,k}_{\mathbf{P}} = \overbrace{Z \times_H Z \times_H \cdots \times_H Z}^{k}$. By the very definition of the fibered product, we get

$$\mathrm{Hom}_{S_0}(S, H^k) = \left\{(h, s_1, \ldots, s_k) \middle| h \in \mathrm{Hom}_{S_0}(S, H),\ s_j \in \mathrm{Hom}_S(S, Z)\right\}.$$

Thus H^k classifies all flat closed subschemes with Hilbert polynomial P having k sections over S. The projection $p : Z^{(k)} = Z \times_H H^k \to H^k$ ($p(z_0 \times z) = z$) gives the universal scheme and has tautological k sections

$$\sigma_j : H^k \ni (z) \mapsto (z_j \times z) \in Z^{(k)} \quad (z = (z_1, \ldots, z_k)).$$

This shows that $\mathcal{H}_{d,N} \subset \underline{H}^k$ for $k = 2n + 1$. For simplicity, write H_0 for H^k. Since "smoothness" is an open condition (because it is local; in other words,

smoothness at a point x of a morphism f follows from formal smoothness of the local ring at x over the local ring at $f(x)$; see [NMD] Section 2.2), there is an open subscheme H_1 of H_0 over which Z is smooth. Then H_1 represents smooth closed subschemes in $\mathbf{P}$ with Hilbert polynomial P and k sections.

Now we use a result of Grothendieck. Abelian varieties have rigidity such that in a smooth projective family $X \to S$ for connected locally Noetherian S with a section $\mathbf{0} : S \to X$, if one fiber is an abelian variety with the identity section induced by $\mathbf{0}$, X itself is an abelian scheme ([GIT] Theorem 6.14). This shows that H_1 has a closed subscheme H_2 over which $Z_2 = Z^{(k)} \times_{H_1} H_2$ is an abelian scheme with the identity section $\mathbf{0}$ induced by σ_0.

Rewrite τ_j $(j = 1, \ldots, 2n)$ for the universal $2n$ sections of Z_2 different from $\mathbf{0}$. We have a maximal closed subscheme $H_3 \subset H_2$ defined by the equations $[N] \circ \tau_i = \mathbf{0}$ for all i, where $[N]$ is the multiplication by the integer N.

The relation $\sum_{j=1}^{2n} a_j \tau_j = \mathbf{0}$ for a given $a = (a_j) \in (\mathbb{Z}/N\mathbb{Z})^{2n} - \{0\}$ gives a closed subscheme H_a of H_3; so, we define $H_4 = H_3 - \bigcup_a H_a$. Thus the abelian scheme Z_4 over H_4 has $2n$ linearly independent sections of order N.

Since Z_4 is a subscheme of $\mathbf{P}_{/H_4}$, it has the line bundle $L = \mathcal{O}_{Z_4}(1)$ which is the restriction of $\mathcal{O}(1)_{/\mathbf{P}}$. Then we define H_5 to be the maximal subscheme of H_4 such that $p^*L \cong L^{\Delta}(\lambda)^3$ for a polarization $\lambda : Z_4 \to {}^tZ_4$, where $p : Z_5 = Z_4 \times_{H_4} H_5 \hookrightarrow Z_4$ is the inclusion. In other words, for any morphism $\phi : T \to H_4$, if $\phi^*L \cong L^{\Delta}(\lambda)^3$ for a polarization $\lambda : {}^tZ_{4,T} \to Z_{4,T} = Z_4 \times_{H_4} T$, ϕ has an image in H_5 (i.e., ϕ factors through H_5). It is proved in [GIT] Proposition 6.11 that the maximal subscheme H_5 with the above property exists and is closed in H_4.

6.3.5 Conclusion

By the argument in the previous proposition, the functor $\mathcal{A}'_{d,N}$ is represented over S_0 by a quasi-projective scheme H_5 with the universal abelian scheme Z_5 over H_5. The group $PGL(m+1)$ acts on H_5 by $\iota \mapsto \iota \circ g$ $(g \in PGL(m+1))$. By the following result of Mumford, if N is sufficiently large, we can make the geometric quotient by Theorem 6.19,

$$\mathfrak{M}_{d,N} = PGL(m+1)\backslash H_5 \quad \text{and} \quad \mathbf{X}_{d,N} = PGL(m+1)\backslash Z_5,$$

which gives the universal abelian scheme over $\mathfrak{M}_{d,N}$.

Proposition 6.23 (Mumford) *Let $A_{/k}$ with projective embedding $\phi : A \hookrightarrow \mathbf{P}^m_{/k}$ be an abelian variety of dimension n over an algebraically closed field k. Write g for the degree of $\phi(A) \subset \mathbf{P}^m$. Let N be a positive integer. Arrange all points of order N of $\phi(A)$ as $f_N(A) = (x_1, \ldots, x_{N^{2n}})$ and regard it as a point $f_N(A) = (\phi(x_1), \ldots, \phi(x_{N^{2n}}))$ of $(\mathbf{P}^m)^{N^{2n}}(k)$. Suppose that the image of A in $\mathbf{P}^m$ is not contained in any hyperplane. If $N > \sqrt{(m+1)g}$ with $char(k) \nmid N$, the point $f_N(A) \in (\mathbf{P}^m)^{N^{2n}}$ is stable.*

We repeat here the proof of Mumford in [GIT] Proposition 7.7.

Proof. We need to check the inequality of Definition 6.15 for the point $f_N(A)$. Since the total number of the points is $M = N^{2n}$, we only need to prove

$$\frac{|\{x_1, \ldots, x_M\} \cap H|}{M} < \frac{1}{m+1}$$

for any hyperplane $H \subset \mathbf{P}^m$. Write (Y, Z) for the intersection number of complementary cycles Z and Y in $\mathbf{P}^m$ ([ALG] I.7 and Appendix A). Choose a 1-dimensional cycle $\gamma = \phi(A) \times_{\mathbf{P}^m} L$ for a linear subspace $L \subset \mathbf{P}^m$ of codimension $n-1$. We suppose that γ contains the origin of $\phi(A)$. Since multiplication $[N]$ by N acts on divisors by N^2 ([ABV] Section 6, Corollary 3), it also acts on the 1-cycles by N^2 (by the Albanese–Picard duality; [ABV] Section 23); so, $[N]^*\gamma = N^2\gamma$ up to numerical equivalence. Here we use the fact that $char(k) \nmid N$. Writing $h = \phi^{-1}(H) = \phi(A) \times_{\mathbf{P}^m} H$, we have

$$\begin{aligned} N^2([N]^*\gamma, h) &= ([N]^*\gamma, [N]^*h) = [N]^*(\gamma, h) \\ &= \deg([N])(\gamma, h) = M(\gamma, h) = M \deg(\phi(A)) = M \cdot g. \end{aligned}$$

This shows

$$\frac{|\{x_1, \ldots, x_M\} \cap H|}{M} \leq \frac{g}{N^2} < \frac{1}{m+1},$$

which finishes the proof. □

Out of $\mathbf{0}, \tau_1, \ldots, \tau_{2n}$, we create the set of full sections $s_1, \ldots, s_M : H_5 \hookrightarrow Z_5$ through linear combinations of τ_j. We embed H_5 into $(\mathbf{P}^m)^M$ by sending $x \in H_5$ to $(s_1(x), \ldots, s_M(x)) \in (\mathbf{P}^m)^M$. Since the image is contained in the open set of stable points $U_{st} \subset (\mathbf{P}^m)^M$ fiber-by-fiber; so, it is contained in U_{st}. Since U_{st} is a $PGL(m+1)$-torsor over its geometric quotient (Z, π) in Theorem 6.19, we find the geometric quotient of H_5 by $PGL(m+1)$ as the image $\pi(H_5)$ of H_5 in Z. Since we know the Hilbert polynomial of Z_5,

$$P(T) = 6^n d \cdot T^n = \frac{\deg(Z_5) T^n}{n!},$$

we know the degree $\deg(Z_5)$ (cf. [ALG] I.7.6). Thus if N is sufficiently large as in Proposition 6.23, the quotient $\mathfrak{M}_{d,N} = PGL(m+1)\backslash H_5$ represents the functor $\mathcal{A}_{d,N}$ over $\mathbb{Z}[\frac{1}{N}]$ (see [GIT] Section 3 in Chapter 7 for more details).

Suppose now that $N \geq 3$. We choose two distinct large primes p and q outside dN so that $\mathcal{A}_{d,pN}$ and $\mathcal{A}_{d,qN}$ are representable by the above argument, respectively, over $\mathbb{Z}[\frac{1}{pN}]$ and $\mathbb{Z}[\frac{1}{qN}]$. The finite (constant) group

$$GL_{2n}(\mathbb{Z}/pN\mathbb{Z}) = GL_{2n}(\mathbb{Z}/p\mathbb{Z}) \times GL_{2n}(\mathbb{Z}/N\mathbb{Z})$$

acts on (A, λ, ϕ_{pN}) by $\phi_{pN} \mapsto \phi_{pN} \circ g$ for $g \in GL_{2n}(\mathbb{Z}/pN\mathbb{Z})$. Since we do not impose (A4) just after Theorem 6.20, the full group $GL_{2n}(\mathbb{Z}/pN\mathbb{Z})$ acts on the functor (and hence on the moduli) in place of the symplectic group $Sp_{2n}(\mathbb{Z}/pN\mathbb{Z})$. Since $\mathrm{Aut}(A, \lambda, \phi_N) = \{1\}$ if $N \geq 3$ by a result of Serre (see

[ABV] Theorem 5 in Section 21), the action of the finite group $GL_{2n}(\mathbb{Z}/p\mathbb{Z})$ is free. If we can cover $\mathcal{A}_{d,pN}$ by affine open subschemes stable under the action, we can again make the geometric quotient

$$\mathfrak{M}_{d,N/\mathbb{Z}[\frac{1}{pN}]} = GL_{2n}(\mathbb{Z}/p\mathbb{Z})\backslash\mathfrak{M}_{d,pN/\mathbb{Z}[\frac{1}{pN}]},$$

which represents $\mathcal{A}_{d,N}$ over $\mathbb{Z}[\frac{1}{pN}]$ (see [SGA] V.1, [ABV] Section 12, or [GME] 1.8.3). In order to give a sketch of a proof of finding an affine open cover of $\mathfrak{M} := \mathfrak{M}_{d,pN}$ stable under the action, we consider the universal abelian scheme $\mathbf{X} \xrightarrow{\pi} \mathfrak{M}$ and $\underline{\omega} = \det(\pi_*\Omega_{\mathbf{X}/\mathfrak{M}})$. The line bundle $\underline{\omega}$ is ample (cf. [CSM] V.3, or see the following subsection for another proof using a compactification of the moduli). By the existence of nontrivial sections θ invariant under $GL_{2n}(\mathbb{Z}/p\mathbb{Z})$ (given by theta constants) of $\underline{\omega}^{\otimes k}$ for sufficiently large k, $\mathfrak{M}[\frac{1}{\theta}]$ is affine and stable under $GL_{2n}(\mathbb{Z}/p\mathbb{Z})$. Since abelian schemes are embedded into a projective space by algebraic theta functions giving rise to theta constants of level N if $N \geq 3$ (see [CSM] Appendix II), moving around θ, $\mathfrak{M}[\frac{1}{\theta}]$ covers $\mathfrak{M}$. Thus we have the geometric quotient $\mathfrak{M}_{d,N/\mathbb{Z}[\frac{1}{pN}]}$ as above.

By the universality, the two schemes $\mathfrak{M}_{d,N/\mathbb{Z}[\frac{1}{pN}]}$ and $\mathfrak{M}_{d,N/\mathbb{Z}[\frac{1}{qN}]}$ thus obtained glue together uniquely over $\mathbb{Z}[\frac{1}{pqN}]$ giving rise to the scheme $\mathfrak{M}_{d,N/\mathbb{Z}[\frac{1}{N}]}$ representing the functor $\mathcal{A}_{d,N}$ over $\mathbb{Z}[\frac{1}{N}]$. Even if $N \leq 3$, the above process yields the coarse moduli scheme $\mathfrak{M}_{d,N}$ over $\mathbb{Z}[\frac{1}{N}]$ (cf. [CSM] Theorem 1.4, or [GME] 2.3.2 in the elliptic modular case).

Once representability of $\mathcal{A}_{d,N}$ is shown, as long as $\mathrm{Aut}((A,\phi_N,\lambda)_{/\overline{\mathbb{F}}_p}) = \{1\}$ for any prime $p \nmid dN$, the formal completion of $\mathfrak{M}_{d,N}$ at the closed point $x \in \mathfrak{M}_{d,N}(\overline{\mathbb{F}}_p)$ carrying $(A,\phi_N,\lambda)_{/\overline{\mathbb{F}}_p}$ is isomorphic to a power series ring $W[[t_{ij}]]_{1\leq i\leq j\leq n}$ for the ring W of Witt vectors with coefficients in $\overline{\mathbb{F}}_p$. This follows from the deformation theory of abelian schemes (see [DAV] I.3 for a summary, [CBT] Chapter V for details, and also Section Section 8.2 particularly 8.2.4 in the text). Thus $\mathfrak{M}_{d,N}$ for $N \geq 3$ is smooth over $\mathbb{Z}[\frac{1}{dN}]$. This finishes the proof of Theorem 6.20.

6.3.6 Smooth Toroidal Compactification

Let us describe briefly without proof the smooth toroidal compactification $M_N = M_{1,N}$ of $\mathfrak{M}^J_{N/\mathbb{Z}[\zeta_N,\frac{1}{N}]} = \mathfrak{M}^J_{1,N}$ and the universal semi-abelian scheme $\mathcal{G}_{/M_N}$, following the book by Faltings and Chai [DAV]. We expect the n-dimensional universal abelian scheme $\mathcal{X}_{n/\mathfrak{M}^J_N}$ to degenerate over the gth stratum S_g $(0 < g \leq n)$ of the cuspidal divisor to a semi-abelian scheme $\mathcal{G}$ with an exact sequence $0 \to \mathbb{G}^g_m \to \mathcal{G} \to \mathcal{X}_{n-g} \to 0$. Thus we need to create S_g from the moduli of extensions $0 \to \mathbb{G}^g_{m/S} \to \mathcal{G} \to A_{n-g/S} \to 0$ for an abelian scheme $A_{n-g/S}$ of dimension $n-g$.

In the Hilbert modular case, the situation is simpler, and the universal AVRM degenerates into $\mathbb{G}_m \otimes \mathfrak{a}^*$ at the cuspidal divisor over the cusp $(\mathfrak{a},\mathfrak{b})$. The character group of $\mathbb{G}_m \otimes \mathfrak{a}^*$ is canonically isomorphic to $\mathfrak{a}$ by

$$X(\mathbb{G}_m \otimes \mathfrak{a}^*) = \mathrm{Hom}(\mathbb{G}_m, \mathbb{G}_m) \otimes_{\mathbb{Z}} \mathrm{Hom}_{\mathbb{Z}}(\mathfrak{a}^*, \mathbb{Z}) \cong \mathfrak{a}$$

via the trace pairing $(a, b) \mapsto \mathrm{Tr}_{F/\mathbb{Q}}(ab)$, and $F_\infty = \mathfrak{a} \otimes_{\mathbb{Z}} \mathbb{R}$. Indeed we started our work in Section 4.1.4 with a simplicial cone decomposition of $F_{\infty+}^{\times}$ on which $(\cdot, \cdot)$ is positive-definite. The term q^ξ in the q-expansion is identified with the linear form on $(\mathfrak{a}\mathfrak{b})^*$; so, $\xi \in \mathfrak{a}\mathfrak{b}$.

By an analogy to $X(\mathbb{G}_m \otimes \mathfrak{a}^*)$, we take $X = X(\mathbb{G}_m^g)$ and consider $X_{\mathbb{R}} = X \otimes_{\mathbb{Z}} \mathbb{R}$ (which is the analogue of F_∞). For the Siegel modular forms (of level 1), each term of their Fourier expansion at the infinity cusp is given by $\mathbf{e}(\mathrm{Tr}(\xi z))$ for $z \in \mathfrak{H}_n$ with symmetric semi-integral matrices ξ in $\mathrm{End}(X)$. We identify the space of such matrices with the symmetric second tensor product $S(X) = Sym^2(X)$ ($S(X)$ is the degree-2 component of $\mathbf{S}(X)$ in Section 6.1.1). The $\mathbb{Z}$-dual $S^*(X)$ is the space of $\mathbb{Z}$-valued symmetric bilinear forms on X. Writing X_R for $X \otimes_{\mathbb{Z}} R$, we similarly define $S^*(X_R)$ and $S(X_R)$. In particular, $S^*(X_{\mathbb{R}})$ is the space of symmetric $\mathbb{R}$-bilinear forms on $X_{\mathbb{R}}$, and $S(X_{\mathbb{R}}) = Sym^2(X_{\mathbb{R}})$. We take the cone $C(X_{\mathbb{R}}) \subset S^*(X_{\mathbb{R}})$ made up of all positive semi-definite bilinear forms. The interior $C^\circ(X_{\mathbb{R}})$ is the space of all positive-definite bilinear forms on $X_{\mathbb{R}}$.

For the moment, we assume that $g = n$ (with the maximal toric part). We take a cone decomposition $\mathcal{C} = \{\sigma\}_\sigma$ of $C(X_{\mathbb{R}})$ satisfying the following conditions similar to (PC1–4):

(Pc1) σ is an open simplicial cone ($\sigma = \{0\}$ is allowed);
(Pc2) The cones in $\mathcal{C}$ are permuted under the action of $GL(X)$, and $\mathcal{C}/GL(X)$ is finite, where $\gamma \in GL(X)$ acts on $b(x, y) \in S^*(X_{\mathbb{R}})$ by $b(x, y) \mapsto b(\gamma x, \gamma y)$;
(Pc3) σ is smooth (i.e., generated by a part of a $\mathbb{Z}$-base of $S^*(X)$);
(Pc4) $\{\sigma\}$ is projective, that is, it is sufficiently fine so that the toroidal compactification is projective (see [DAV] V.5.1 for an exact condition for projectivity).

We reproduce here some results of the more detailed theorems in [DAV] IV.6.7, V.2.3 and V.5.8 on the smooth toroidal compactification we need later.

Theorem 6.24 (Chai–Faltings) *Suppose $n > 1$ and $N \geq 3$. There exist projective normal schemes $M_N = M_N^{\mathcal{C}}$ and M_N^* over $\mathbb{Z}[\frac{1}{N}, \zeta_N]$ containing $\mathfrak{M}_N^J$ as an open dense subscheme satisfying the following five conditions.*

(1) *The natural action of $Sp_{2n}(\mathbb{Z})$ on $\mathfrak{M}_N^J$ extends to M_N and M_N^*. The complement $D_\infty = M_N - \mathfrak{M}_N^J$ is a relative Cartier divisor with normal crossings. The scheme M_N is smooth over $\mathbb{Z}[\frac{1}{N}, \zeta_N]$.*
(2) *For $N|N'$, the natural projection $\mathfrak{M}_{N'}^J \twoheadrightarrow \mathfrak{M}_N^J$ extends to $M_{N'} \to M_N$ (resp. $M_{N'}^* \to M_N^*$), which is equivariant under $Sp_{2n}(\mathbb{Z}/N'\mathbb{Z})$.*
(3) *The universal abelian scheme $\mathcal{X}_N$ over $\mathfrak{M}_N^J$ extends to a semi-abelian scheme $\mathcal{G}_N \xrightarrow{\pi} M_N$, and $\mathcal{G}_{N'}$ is the pullback of $\mathcal{G}_N$ if $N|N'$. The Kodaira–Spencer map induces an isomorphism of $Sym^2(\pi_*(\Omega_{\mathcal{G}_N/\mathbb{Z}[\frac{1}{N},\zeta_N]}))$ onto $\Omega_{M_N/\mathbb{Z}[\frac{1}{N},\zeta_N]}[d\log(D_\infty)]$ as locally free sheaves.*

(4) *Let* $\underline{\omega} = \pi_* \Omega_{\mathcal{G}_N/\mathbb{Z}[\frac{1}{N},\zeta_N]}$. *Then* $H^0(M_N, \det(\underline{\omega})^{\otimes_J}) \cong H^0(\mathfrak{M}_N^J, \det(\underline{\omega})^{\otimes_J})$ *by the restriction map (Koecher principle), which vanishes if* $j < 0$.

(5) *Set* $\mathfrak{G}_N = \bigoplus_{j \geq 0} H^0(M_N, \det(\underline{\omega})^{\otimes_J})$. *Then* $\mathfrak{G}_N$ *is a graded algebra finitely generated over* $\mathbb{Z}$, *and we have* $M_N^* \cong Proj(\mathfrak{G}_N)$. *In particular,* M_N *covers* M_N^*, *and* M_N *is the normalization of a blowing up of* M_N^* *along a coherent sheaf of ideals.*

The theorem is proven first constructing M_N as an algebraic stack using only (Pc1–3). Then using (Pc4), an ample line bundle is created on the algebraic stack M_N, and hence M_N is a projective scheme. Even if $N \leq 2$, the algebraic stack M_N (and the semi-abelian stack $\mathcal{G}_N$ over M_N) satisfies the above condition of theorems (except for the smoothness), and (the reduced scheme of) each stratum of the cuspidal divisor D_∞ is smooth over $\mathbb{Z}[\frac{1}{N}, \zeta_N]$. The expression $M_N^* = Proj(\mathfrak{G}_N)$ is valid for $N \leq 2$, and M_N^* is always a normal scheme. The scheme M_N^* is called the *minimal* compactification (with $M_N^*(\mathbb{C})$ giving the topological Satake compactification) and M_N is called the smooth projective *toroidal* compactification.

Here is a very brief description of M_N and some indication of its construction. Since we need the modulus variety $\mathfrak{M}_{1,N}^J$ for g-dimensional (principally polarized) abelian schemes for $g = 1, \ldots, n$ in order to describe M_N, we write $\mathfrak{M}_N^{(g)}$ for the level N modulus $\mathfrak{M}_N^J$ of principally polarized abelian schemes of dimension g (dropping the superscript J for simplicity). The reader is assumed to be familiar with the notation of [DAV].

We start with the classification theory of Raynaud extensions (to see what are the moduli of extensions $0 \to \mathbb{G}_m^g \to G \to A_{n-g} \to 0$). Since the algebraic group G admitting its Shimura variety does not appear in this subsection, we use the symbol G (following [DAV]) to indicate a Raynaud extension (this convention is in effect only in this subsection; the symbol $\mathcal{G}$ is used to indicate the quotient semi-abelian scheme of G by a unit action). Let $S = Spec(R)$ for an I-adically complete Noetherian integral domain $R = \varprojlim_k R/I^k$. We put $S_0 = Spec(R/I)$. Start with an extension $T_0 \hookrightarrow G_0 \twoheadrightarrow A_0$ over S_0 of an abelian scheme A_{0/S_0} by a torus T_{0/S_0}. Pick a semi-abelian scheme $G_{/S}$ with reduction modulo I giving rise to $T_0 \hookrightarrow G_0 \twoheadrightarrow A_0$. Writing $S_\infty = Spf(R)$ for the formal completion of S along S_0, it is known (cf. [NMD] 7.3–4 and [CSM] Chapter II) that T_0 is uniquely lifted to a formal torus $\widehat{T} \subset \widehat{G}$, where $\widehat{G}$ is the formal completion of G along G_0. The quotient $\widehat{A} = \widehat{G}/\widehat{T}$ is a smooth formal group, which is proper because otherwise A_0 has to have an affine subgroup scheme. Thus $\widehat{A} \mod I^j$ for all $j > 0$ is an abelian scheme over $Spec(R/I^j) = S_j$. We call such a formal scheme a *formal abelian scheme*. When there exists a subtorus $T_{/S} \subset G_{/S}$ such that the quotient $A_{/S} = G/T$ is an abelian scheme and the extension $T \hookrightarrow G \twoheadrightarrow A$ gives rise to the formal extension $\widehat{T} \hookrightarrow \widehat{G} \twoheadrightarrow \widehat{A}$, we call $T \hookrightarrow G \twoheadrightarrow A$ a *Raynaud extension*, which is uniquely determined by $\widehat{T} \hookrightarrow \widehat{G} \twoheadrightarrow \widehat{A}$ (and hence by $G_{/S}$) by the formal existence theorem of Grothendieck ([EGA] III.5.1.4 applied to A). If $\widehat{G}$ is

quasi-projective having an ample invertible sheaf $\mathcal{L}$ on $\widehat{G}$, then the $\widehat{T}$-fixed part of $\mathcal{L} \otimes [-1]^*\mathcal{L}$ descends to an ample sheaf on $\widehat{A}$ ([DAV] II.1 Proposition 1.1); so, $\widehat{A}$ is projective. Again by the formal existence theorem of Grothendieck ([EGA] III 5.4.5), $\widehat{A}$ is algebraizable. Then $\mathrm{Ker}(G \twoheadrightarrow A)$ has to be a torus $T_{/S}$ since its formal completion along T_0 is a formal torus. In our application, R is a formal monoid ring $\mathbb{Z}[[q^\xi]]_\xi$ (i.e., the q-expansion ring) where ξ runs over $\sigma^\vee \cap \frac{1}{N}S(X)$ for the dual cone $\sigma^\vee$ of $\sigma \in \mathcal{C}$, where $X = X(\mathbb{G}_m^g)$.

Starting from a split torus $T_{/S}$ and an abelian scheme $A_{/S}$, we classify all Raynaud extensions of A by T in terms of the étale sheaf $\underline{X}(V) = \mathrm{Hom}_{\text{alg-gp}}(T_V, \mathbb{G}_{m/V})$ for S-scheme V and $T_V = T \times_S V$. By definition, $\underline{X}$ is a constant torsion-free sheaf of $\mathbb{Z}$-rank g ($\Leftrightarrow \dim T = g$) under the étale topology over S, because T is split. Let $A_{/S}$ be an abelian scheme of relative dimension $n-g$. Since $\underline{X}$ is constant, we often regard $\underline{X}$ just as a $\mathbb{Z}$-free module of rank g and denote it simply by X. Let $Y = \mathrm{Hom}_{\mathbb{Z}}(X, \mathbb{Z}) = X^*$. Define a split torus T over S by

$$T = \mathrm{Hom}_{\mathbb{Z}}(X, \mathbb{G}_m) = \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}, X^* \otimes_{\mathbb{Z}} \mathbb{G}_m) = Y \otimes_{\mathbb{Z}} \mathbb{G}_m.$$

Then we have

$$\begin{aligned}\mathrm{Hom}_{S\text{-alg-gp}}(T, \mathbb{G}_m) &= \mathrm{Hom}_{S\text{-alg-gp}}(Y \otimes_{\mathbb{Z}} \mathbb{G}_m, \mathbb{G}_m) \\ &\cong (X^*)^* \otimes \mathrm{Hom}_{S\text{-alg-gp}}(\mathbb{G}_m, \mathbb{G}_m) = X.\end{aligned}$$

For any given extension over an S-scheme S',

$$0 \to T_{/S'} = T \times_S S' \to G_{/S'} \xrightarrow{\pi} A_{/S'} = A \times_S S' \to 0,$$

we have the following exact sequence,

$$\mathrm{Hom}_{S\text{-alg-gp}}(T, \mathbb{G}_m) \xrightarrow{\phi_G} \mathrm{Ext}^1_{S\text{-alg-gp}}(A, \mathbb{G}_m) \xrightarrow{\pi^*} \mathrm{Ext}^1(G, \mathbb{G}_m).$$

Note that $\mathrm{Hom}_{S\text{-alg-gp}}(T, \mathbb{G}_m) = X_{/S}$ and that the dual abelian scheme ${}^tA_{/S}$ is equal to $\mathrm{Ext}^1_{S\text{-alg-gp}}(A, \mathbb{G}_m)$. See [CSM] Remark 2.1.4, group extensions of A by $\mathbb{G}_m$ (not just scheme extensions) correspond to $\mathbb{G}_m$-torsors (hence line bundles) in the identity-connected component of $\mathrm{Pic}_{A/S}$. Thus we have an association $G \mapsto \phi_G \in \mathrm{Hom}_{\mathbb{Z}}(X, {}^tA)$.

On the other hand, we take the universal extension

$$0 \to \mathbb{G}_m \to G \to A \to 0$$

over ${}^tA \subset \mathrm{Pic}_{A/S}$. We fix a base $\xi_1, \ldots, \xi_g$ of X. Then consider the morphism

$$T = \mathrm{Hom}_{\mathbb{Z}}(X, \mathbb{G}_m) \to \mathbb{G}_m$$

given by $\phi \mapsto \sum_i \phi(\xi_i)$. Similarly, we define $\mathrm{Hom}_{\mathbb{Z}}(X, {}^tA) \to {}^tA$ by using the same base. Thus we have a commutative square

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbb{Z}}(X,\mathbb{G}_m) = T & \longrightarrow & \mathbb{G}_m \\ \downarrow & & \downarrow \\ \mathrm{Hom}_{\mathbb{Z}}(X,{}^tA) & \longrightarrow & {}^tA\,. \end{array}$$

Pulling back the above exact sequence $0 \to \mathbb{G}_m \to G \to A \to 0$ over ${}^tA$ by this square, we get $0 \to \mathrm{Hom}_{\mathbb{Z}}(X,\mathbb{G}_m) = T \to \widetilde{G} \to A \to 0$ over $\mathrm{Hom}_{\mathbb{Z}}(X,{}^tA)$ such that for any given $\chi \in X$, the morphism $\mathrm{Hom}_{\mathbb{Z}}(X,{}^tA) \to {}^tA$ given by $\phi \mapsto \phi(\chi)$ gives rise (by push-forward) to an extension

$$0 \to \mathbb{G}_m = T/\operatorname{Ker}(\chi) \to \widetilde{G}/\operatorname{Ker}(\chi) \to A \to 0$$

which is associated with $\phi(\chi) \in {}^tA$. Thus the pair $(\mathrm{Hom}_{\mathbb{Z}}(X,{}^tA_{/S}),\widetilde{G})$ represents the functor $S' \mapsto \left[T_{/S'} \hookrightarrow G \twoheadrightarrow A_{/S'}\right]$ on the category of schemes over S (see [CSM] 2.1–2 for more details).

For a given Raynaud extension $T \hookrightarrow G \twoheadrightarrow A$ over S given by $\phi_G \in \mathrm{Hom}(X,{}^tA)$, we can construct its dual ${}^tG = G/K(\mathcal{L})$ (in the same manner as the construction of ${}^tA$ by Mumford [ABV] Section 13; see [DAV] II.2). Here $K(\mathcal{L})$ is the stabilizer of $\mathcal{L}$ under the translation action by G. The dual ${}^tG$ is again a Raynaud extension ${}^tT \hookrightarrow {}^tG \twoheadrightarrow {}^tA$ for the dual torus ${}^tT = \mathrm{Hom}(Y,\mathbb{G}_m) = X \otimes_{\mathbb{Z}} \mathbb{G}_m$ ($\Leftrightarrow X({}^tT) = Y = \mathrm{Hom}_{\mathbb{Z}}(X,\mathbb{Z})$). Thus we have $\phi_{{}^tG} \in \mathrm{Hom}(Y,A)$ corresponding to this dual Raynaud extension.

We can generalize the above construction with respect to (R,I) to an integral scheme S (not necessarily affine) and a sheaf of ideals $\mathfrak{I}$, because the universal schemes $\mathrm{Hom}(X,{}^tA)_{/O}$ for affine open subsets $O = Spec(R) \subset S$ glue together by universality.

We now assume $X = X(\mathbb{G}_m^n)$; so, $C(X_{\mathbb{R}}) = \bigsqcup_{\sigma\in\mathcal{C}} \sigma$. We write X_σ for $X/\cap_{b\in\sigma} Rad(b)$ for the radical $Rad(b)$ (i.e., the null space) of b. The space $S(X_{\sigma,\mathbb{R}})$ is a quotient of $S(X_{\mathbb{R}})$, and by duality we have $S^*(X_{\sigma,\mathbb{R}}) \subset S^*(X_{\mathbb{R}})$, where $X_{\sigma,\mathbb{R}} = X_\sigma \otimes_{\mathbb{Z}} \mathbb{R}$. We write ξ for the collection of σ giving the same $\cap_{b\in\sigma} Rad(b)$ and write $X_\xi = X_\sigma$ for $\sigma \in \xi$ (which is independent of $\sigma \in \xi$). Then $\mathcal{C}_\xi = \{\sigma \in \mathcal{C} | \sigma \subset S^*(X_{\xi,\mathbb{R}})\}$ is a cone decomposition of $X_{\xi,\mathbb{R}}$ satisfying (Pc1–4). Let $g = \operatorname{rank}_{\mathbb{Z}} X_\sigma$. Then each boundary stratum $\mathcal{S}_\sigma$ (corresponding σ of dimension g) is constructed from a universal extension ([DAV] IV.3) over $Z_\sigma = \mathrm{Hom}_{\mathbb{Z}}(X_\sigma,{}^t\mathcal{X}_\sigma)$: $\mathrm{Hom}_{\mathbb{Z}}(X_\sigma,\mathbb{G}_m) \hookrightarrow \widetilde{G}_\sigma \twoheadrightarrow \mathcal{X}_\sigma$, where $\mathcal{X}_\sigma = \mathcal{X}_{n-g}$ ($g = \operatorname{rank} X_\sigma$) is the $(n-g)$-dimensional principally polarized universal abelian scheme. The scheme $\mathcal{S}_\sigma$ roughly corresponds to $Spec(R_\sigma)$ in Section 4.1.4 in the Hilbert modular case but is more complicated because of the contribution of the abelian variety quotient $\mathcal{X}_\sigma$. The group $GL(X)$ acts on $\bigcup_\sigma \mathcal{S}_\sigma$ through X_σ (and permuting $\sigma \in \mathcal{C}$). Roughly speaking, we can canonically glue $\{\mathcal{S}_\sigma\}_{\sigma\in\mathcal{C}}$ into a scheme $\mathcal{S} = \bigcup_\sigma \mathcal{S}_\sigma$ as in the Hilbert modular case, and the quotient of the formal completion along the boundary of $\mathcal{S}$ by $GL(X)$ gives the formal completion of M_1 along the cuspidal divisor D_∞. Strictly speaking, first we make the quotient $\bigcup_{\sigma\in\mathcal{C}_\xi} \mathcal{S}_\sigma$ by $GL(X_\sigma)$, and then we again glue these schemes by the action of $GL(X)$.

We give an indication of the construction of $\mathcal{S}_\sigma$ and the semi-abelian scheme $\mathcal{G}_\sigma$ over $\mathcal{S}_\sigma$ with a symplectic level N-structure for $N \geq 3$ (when M_N is a scheme). Hereafter N is invertible on the base S, and suppose for the moment that we have a symplectic level N-structure $\alpha : \mathcal{G}[N] \cong (\mathbb{Z}/N\mathbb{Z})^{2n}$. Since $\mathcal{G} = \text{“}\widetilde{G}/q(Y)\text{”}$, we have an exact sequence: $0 \to \widetilde{G}[N] \to \mathcal{G}[N] \to Y_N/Y \to 0$, where $Y_N = \frac{1}{N}Y$. By moving around by an element of $Sp_{2n}(\mathbb{Z}/N\mathbb{Z})$ acting on $(\mathbb{Z}/N\mathbb{Z})^{2n}$, we may assume that α induces an isomorphism $\alpha_Y : Y_N/Y \to (\mathbb{Z}/N\mathbb{Z})^g$ and $\widetilde{\alpha} : \widetilde{G}[N] \cong (\mathbb{Z}/N\mathbb{Z})^{2n-g}$ so that the symplectic pairing on the target induces the duality between $T[N] \cong \mu_N^g$ and Y_N/Y. The stabilizer of $\alpha(\widetilde{G}[N])$ gives rise to a maximal parabolic subgroup $P_{n,g}(\mathbb{Z}/N\mathbb{Z}) \subset Sp_{2n}(\mathbb{Z}/N\mathbb{Z})$, and we can extend $q : Y \to {}^tA$ to $q_N : Y_N \to {}^tA$. Since A is principally polarized, we identify $Y = X$ and ${}^tA = A$.

Now we take $S = \mathfrak{M}^{(n-g)}$ and $\mathcal{X} = \mathcal{X}_{n-g}$ for the universal abelian scheme of dimension $n-g$ with level N-structure. Let $0 \to T_\sigma \to \widetilde{G}_\sigma \to \mathcal{X} \to 0$ be the universal extension over $Z_\sigma = \mathrm{Hom}_{\mathbb{Z}}(X_\sigma, {}^t\mathcal{X})$. We write $E_\sigma = S^*(X_\sigma) \otimes_{\mathbb{Z}} \mathbb{G}_m$, which is the split torus associated with the character group $S(X_\sigma)$. Then $S(X_\sigma) = \mathrm{Hom}_{\mathbb{Z}}(S^*(X_\sigma), \mathbb{Z})$ for the space $S^*(X_\sigma)$ of symmetric bilinear forms $b(x,y) = b(y,x)$. Let $\sigma^\vee$ be the dual cone

$$\sigma^\vee = \{x \in S(X) | y(x) \geq 0 \ \forall y \in \sigma\}.$$

Then we make the torus embedding

$$E_\sigma = Spec(\mathbb{Z}[S(X_\sigma)]) \hookrightarrow \overline{E}_\sigma = Spec(\mathbb{Z}[\sigma^\vee \cap S(X_\sigma)]).$$

Let $\mu, \nu \in X_\sigma$. Then the symmetrization $[\mu \otimes \nu] \in S(X_\sigma)$ of $\mu \otimes \nu \in X_\sigma \otimes X_\sigma$ defines a character $[\mu \otimes \nu] : E_\sigma \to \mathbb{G}_m$. We consider the abelian scheme $\mathcal{X} \times_{\mathfrak{M}^{(n-g)}} {}^t\mathcal{X} = \mathcal{X} \times_{\mathfrak{M}^{(n-g)}} \mathcal{X}$ on $\mathfrak{M}^{(n-g)}$. Here we identify ${}^t\mathcal{X}$ with $\mathcal{X}$ by the universal principal polarization. We have a tautological map

$$c(\mu) : Z_\sigma = \mathrm{Hom}_{\mathbb{Z}}(X_\sigma, {}^t\mathcal{X}) \to {}^t\mathcal{X} = \mathcal{X}$$

sending $\phi : X_\sigma \to {}^t\mathcal{X}$ to $\phi(\mu)$. Recall the Poincaré line bundle $\mathcal{P}$, that is, the universal line bundle on $\mathcal{X} \times {}^t\mathcal{X}/{}^t\mathcal{X}$ such that for each S-valued point $S \xrightarrow{\alpha} {}^t\mathcal{X}$, $\alpha^*\mathcal{P} \in Pic^0(\mathcal{X}_{/S})$ is the line bundle represented by α. Let $\mathcal{P}^\times = \mathcal{P} - \{0\}$ be the $\mathbb{G}_m$-torsor associated with $\mathcal{P}$.

We then consider

$$c([\mu \otimes \nu]) = c(\mu) \times c(\nu) : Z_\sigma = \mathrm{Hom}_{\mathbb{Z}}(X_\sigma, {}^t\mathcal{X}) \to \mathcal{X} \times_{Z_\sigma} \mathcal{X}.$$

We have a $\mathbb{G}_m$-torsor $\mathcal{X}[\mu \otimes \nu] = c([\mu \otimes \nu])^*(\mathcal{P}^{-1})^\times$ over $Z_\sigma = \mathrm{Hom}_{\mathbb{Z}}(X_\sigma, {}^t\mathcal{X})$. Now we choose a base $B = \{[\mu_i \otimes \nu_i] | i = 1, \ldots, s\}$ of $S(X_\sigma)$. Then $\Xi_\sigma = \prod_i \mathcal{X}[\mu_i \otimes \nu_i]$ is an E_σ-torsor, because $E_\sigma = \mathrm{Hom}_{\mathbb{Z}}(S(X_\sigma), \mathbb{G}_m) \cong \mathbb{G}_m^B$ by $\phi \mapsto \prod_i \phi([\mu_i \otimes \nu_i])$. We symbolically write $\Xi_\sigma = S(X_\sigma) \otimes_{\mathbb{Z}, c^*} (\mathcal{P}^{-1})^\times$. The construction of Ξ_σ does not depend on the choice of the base, and the push-forward of Ξ_σ along $[\mu \otimes \nu] : E_\sigma \to \mathbb{G}_m$ is isomorphic to $\mathcal{X}[\mu \otimes \nu]$. This

follows from the symmetry of $\mathcal{P}$ and the fact that $S(X_\sigma) \to \mathrm{Pic}(Z_\sigma)$ given by $n \mapsto c(n)^*\mathcal{P}$ is a group homomorphism. Since Ξ_σ is the product of the pullback of $\mathcal{P}$ itself by $c \times c$, $c(n)^*\mathcal{P}$ has a tautological section (or trivialization) τ over Ξ_σ. This trivializes the biextension associated with $\mathcal{P}$ and hence gives rise to a homomorphism $i : X_\sigma \to \widetilde{G}_\sigma$ lifting the tautological one $c = {}^t c : X_\sigma \to {}^t\mathcal{X} = \mathcal{X}$ on Z_σ over Ξ_σ (see [CSM] II, (2.3.3) for the details and [Mu] for biextensions).

We form the contraction-product $\overline{\Xi}_\sigma = \Xi_\sigma \times^{E_\sigma} \overline{E}_\sigma = (\Xi_\sigma \times_{Z_\sigma} \overline{E}_\sigma)/E_\sigma$ under the diagonal action. This product has the following universality.

1. We have the following commutative diagram

$$\begin{array}{ccc} \overline{E}_\sigma \times_{Z_\sigma} \Xi_\sigma & \xrightarrow{\overline{m}} & \overline{\Xi}_\sigma \\ {\scriptstyle \iota}\uparrow & & \uparrow{\scriptstyle \bar{\imath}} \\ E_\sigma \times_{Z_\sigma} \Xi_\sigma & \xrightarrow[m]{} & \Xi_\sigma; \end{array}$$

2. If we have the following commutative diagram

$$\begin{array}{ccc} \overline{E}_\sigma \times_{Z_\sigma} \Xi_\sigma & \xrightarrow{\overline{\mu}} & F \\ {\scriptstyle \iota}\uparrow & & \uparrow{\scriptstyle \bar{\iota}} \\ E_\sigma \times_{Z_\sigma} \Xi_\sigma & \xrightarrow[\mu]{} & \Xi_\sigma, \end{array}$$

we have a unique morphism $\gamma : \overline{\Xi}_\sigma \to F$ such that $\gamma \circ \overline{m} = \overline{\mu}$ and $\gamma \circ \bar{\iota} = \bar{i}$.

We have a polyhedral cone decomposition $C(X_{\xi,\mathbb{R}}) = \bigsqcup_{\sigma \in \mathcal{C}_\xi} \sigma$. Here we recall that $C(X_{\xi,\mathbb{R}}) \subset S^*(X_{\xi,\mathbb{R}})$ for $X_{\xi,\mathbb{R}} = X_\xi \otimes_{\mathbb{Z}} \mathbb{R}$ is the cone made up of positive semi-definite symmetric bilinear forms. Then all the data $\overline{\Xi}_\sigma$, $\overline{E}_\sigma$, and so on, so far constructed are functorial and glue together over σs in $\mathcal{C}_\xi$ yielding $\overline{\Xi}_\xi$ and so on. We write $W_\xi = \bigsqcup_{\sigma \subset C^\circ(X_\xi)} Z(\sigma)$, where $Z(\sigma) = Spec(\mathbb{Z}[q]_{q|_\sigma = 0, q \in S(X_\sigma)})$. The locally closed subscheme $\Xi_\xi \times^{E_\xi} W_\xi$ of $\overline{\Xi}_\xi$ (given by the contraction product) should be the loci of the quasi-abelian scheme with maximal split torus of rank $\geq g$. We write $\mathcal{S}_\xi$ for the formal completion of $\overline{\Xi}_\xi$ along $\Xi_\xi \times^{E_\xi} W_\xi$. Then we have a semi-abelian scheme ${}^\heartsuit G \to \mathcal{S}_\xi/GL(X_\xi)$ as in [DAV] IV.3.1 with the maximal split torus of rank g on $\Xi_\xi \times^{E_\xi} \overline{W}_\xi$ for $g = \mathrm{rank}\, X_\xi$.

We now incorporate the level N-structure. Let $\alpha : \mathcal{G}[N] \cong (\mathbb{Z}/N\mathbb{Z})^{2n}$ be a symplectic level structure. This induces $q_n : Y_{\xi,N} = \frac{1}{N} Y_\xi \to \mathcal{X}$ ($Y_\xi = \mathrm{Hom}_{\mathbb{Z}}(X_\xi, \mathbb{Z})$), which by duality induces $c_n : X_{\xi,N} = \frac{1}{N} X_\xi \to {}^t\mathcal{X}$. This gives rise to an extension $0 \to T_{\xi,N} = \mathrm{Hom}_{\mathbb{Z}}(X_{\xi,N}, \mathbb{G}_m) \to \widetilde{G}_{\xi,N} \to \mathcal{X} \to 0$ over $\mathrm{Hom}_{\mathbb{Z}}(X_{\xi,N}, {}^t\mathcal{X})$. By construction, $q_N : Y_{\xi,N} \to \mathcal{X}$ lifts to $q_N : Y_{\xi,N} \to \widetilde{G}_{\xi,N}$ so that the quotient is a principally polarized abelian scheme generically. Then we have an exact sequence

$$0 \to Y_{\xi,N}/Y \to \widetilde{G}_{\xi,N}/q_N(Y) \to \widetilde{G}_{\xi,N}/q_N(Y_{\xi,N}) \to 0.$$

This semi-abelian scheme $\widetilde{G}_{\xi,N}/q_N(Y_{\xi,N})$ has a polarization

$$\lambda_N : \widetilde{G}_{\xi,N}/q_N(Y_{\xi,N}) \to {}^t(\widetilde{G}_{\xi,N}/q_N(Y_{\xi,N}))$$

of degree N^{2n}. By definition, $\mathcal{G} = \mathcal{G}_N = (\widetilde{G}_N/q_N(Y_{\xi,N}))/T_{\xi,N}[N]$, which acquires a principal polarization (because of degree N^g-isogeny image) and an injection $Y_{\xi,N}/Y_\xi \hookrightarrow \mathcal{G}_\eta[N]$ with an isotropic image. Thus any level N-structure on $\mathcal{X}$ gives rise to a level N-structure of $\mathcal{G}$. As shown in [DAV] IV.6.5, this construction gives all possible level N-structures on $\mathcal{G}$ after translation by $Sp_{2n}(X/NX)$, and the stabilizer of this level N-structure is $P_{n,g}(X/NX)$. Thus the number of cusps is the index of $P_{n,g}(X/NX)$ in $Sp_{2n}(X/NX)$.

We write $GL_n(X)(N)$ for the kernel of the reduction map of $GL(X)$ into the group $GL(X/NX)$ modulo N. Since the integral structure is given by $N \cdot S^*(X)$, accordingly we need to take $T_{\xi,N} = \mathrm{Hom}_{\mathbb{Z}}(X_{\xi,N}, \mathbb{G}_m)$ instead of T_ξ, because the parameterization is given by $0 \to T_{\xi,N} \to \widetilde{G}_N \to \mathcal{X}_{N,n-g} \to 0$ for the universal abelian scheme $\mathcal{X}_{N,n-g}$ with level N-structure. We have the tautological map $c_{\xi,N} : X_{\xi,N} \to {}^t\mathcal{X}_{N,n-g} = \mathcal{X}_{N,n-g}$. The multiplication by N induces $E_\xi = E_{\xi,N} \to E_\xi$, and we get the torus embedding $E_{\xi,N} \hookrightarrow \overline{E}_{\xi,N}$. The character group of $E_{\xi,N}$ is given by $\frac{1}{N}S(X)$. The universal semi-abelian scheme $\mathcal{G}_N$ with level N-structure is given by

$$(\widetilde{G}_N/q_N(Y))/T_{\xi,N}[N] = \mathcal{G}_N$$

over $\mathrm{Ind}_{P_{n,g}}^{Sp(2n)} \mathcal{S}_{\xi,N} \times \mathrm{Isom}(Y_{\xi,N}/NY_{\xi,N}, (\mathbb{Z}/N\mathbb{Z})^g)/GL(Y_\xi)$. Here the formal scheme $\mathrm{Ind}_{P_{n,g}}^{Sp(2n)} \mathcal{S}_{\xi,N}$ is the disjoint union of copies of the scheme $\mathcal{S}_{\xi,N}$ indexed by the equivalence classes of cusps $Sp_{2n}(X/NX)/P_{n,g}(X/NX)$ on which $Sp_{2n}(X/NX)$ acts by permutation of the indices.

Since we have already studied via Tate curves the compactification of $\mathfrak{M}^J = \mathfrak{M}_N^{(n)}$ when $n = 1$ (the modulus of elliptic curves), we assume here $n > 1$. To make $\mathfrak{M}^J$ smooth, we assume that $N \geq 3$. We then define $\underline{\omega} = f_*\Omega_{\mathcal{X}/\mathfrak{M}^J}$ for the universal abelian scheme $f : \mathcal{X}_N \to \mathfrak{M}^J$. This is a locally free sheaf over $\mathfrak{M}^J$ of rank n. We define $\det \underline{\omega} = \bigwedge^n \underline{\omega}$. The vector bundle $\underline{\omega}$ extends to the toroidal compactification M_N by defining $\underline{\omega} = f_*\Omega_{\mathcal{G}_N/M_N}$ for the semi-abelian scheme $\mathcal{G}$ extending $\mathcal{X}$ to M. Then we define a graded algebra

$$\mathfrak{G} = \mathfrak{G}_N = \bigoplus_{m=0}^{\infty} H^0\left(M_N, \det(\underline{\omega})^{\otimes m}\right) \overset{(*)}{=} \bigoplus_{m=0}^{\infty} H^0\left(\mathfrak{M}_N^J, \det(\underline{\omega})^{\otimes m}\right).$$

The last identity $(*)$ follows from Koecher's principle (Theorem 6.24 (4)) if $n > 1$. Choosing the toroidal compactification data well, M_N is projective ([DAV] V.2.5) so that this graded algebra $\mathfrak{G}$ is finitely generated over $\mathbb{Z}[\frac{1}{N}]$, and by the first equality, the graded algebra is normal. Thus we may define $\mathfrak{G}_N$

to be the normalization of $\mathfrak{G}_1$ in the algebra $\mathcal{R}_{d,N}$ defined below. We define the minimal compactification by $M_N^* = Proj(\mathfrak{G}_N)$. It is called "minimal" because any smooth toroidal compactification M_N covers M_N^*.

We can define a sheaf of graded algebras over M_N by

$$\mathcal{R} = \mathcal{R}_N = \bigoplus_{m=-\infty}^{\infty} f_* \left(\det(\underline{\omega})^{\otimes m}\right).$$

Then $\mathcal{M}_N = Spec_M(\mathcal{R})$ represents the following functor

$$\mathcal{P}_N(S) = \left[(A, \phi_N, \lambda, \omega)_{/S} \middle| (A, \phi_N, \lambda) \in \mathcal{A}_{1,N}(S), H^0(A, \det \Omega_{A/S}) = \mathcal{O}_S \omega\right].$$

Then we have $M_N = \mathbb{G}_m \backslash \mathcal{M}_N$ as a geometric quotient, and $\mathcal{M}_N$ is a $\mathbb{G}_m$-torsor if $N \geq 3$. Here $a \in \mathbb{G}_m$ acts on the functor $\mathcal{P}_N$ by $(A, \phi_N, \lambda, \omega) \mapsto (A, \phi_N, \lambda, a\omega)$. The relation between the modulus scheme classifying abelian schemes with level structure and the one classifying with extra information of nowhere-vanishing n-differentials is exactly the same as in the elliptic modular case, and it is also interesting that this is proven only after the hard work of smoothly compactifying the open modulus $\mathfrak{M}_N^J$.

Remark 6.25 Let the notation and the assumption be as in Remark 6.21. Carrying out the toroidal compactification process for $\mathfrak{M}_{d,N}^{J'}$ to obtain a smooth compactification $M_{d,N}^{J'}$ over $\mathbb{Z}[\zeta_N, \frac{1}{dN}]$, we apply Corollary 6.11 to the semi-abelian scheme $\mathcal{G}_{/M_{d,N}^{J'}}$. Then, automatically, we get a toroidal compactification of the moduli $\mathfrak{M}(\mathfrak{c}, \Gamma(N))$ with respect to the data induced from the toroidal compactification data of $S^*(O)$ (as long as they are compatible). Here $\mathfrak{d}^{-1}$ is embedded into $S^*(O)$ regarding $\xi \in \mathfrak{d}^{-1}$ as a symmetric bilinear form on $O \times O$ by $(x, y) \mapsto \mathrm{Tr}_{F/\mathbb{Q}}(\xi xy)$. An important compatibility of the toroidal compactification data is that the polyhedral decomposition of positive semi-definite $C(O_{\mathbb{R}}) \subset S^*(O_{\mathbb{R}})$ has to induce that on $F_{\infty+}^{\times} \subset \mathfrak{d}^{-1} \otimes_{\mathbb{Z}} \mathbb{R} = F_\infty$ by taking intersection with the image of $\mathfrak{d}^{-1}$.

6.4 Siegel Modular Variety

In this section, we study the Shimura variety associated with $G = GSp(2n)_{/\mathbb{Q}}$, where, for a ring R,

$$GSp_{2n}(R) = \left\{\alpha \in GL_{2n}(R) \middle| {}^t\alpha J_n \alpha = \nu(\alpha) J_n \ \text{ with } \nu(\alpha) \in R^\times\right\}$$

for $J_n = \left(\begin{smallmatrix} 0 & -1_n \\ 1_n & 0 \end{smallmatrix}\right)$ with $0 < n \in \mathbb{Z}$. Consider a homomorphism $h_{\mathbf{0}} : \mathbb{S}_{/\mathbb{R}} := \mathrm{Res}_{\mathbb{C}/\mathbb{R}} \mathbb{G}_m \to GSp(2n)_{/\mathbb{R}}$ of algebraic groups given by $\mathbb{S}(\mathbb{R}) = \mathbb{C}^\times \ni a + bi \mapsto \left(\begin{smallmatrix} a \cdot 1_n & b \cdot 1_n \\ -b \cdot 1_n & a \cdot 1_n \end{smallmatrix}\right) \in GSp_{2n}(\mathbb{R})$. Let X be the collection of all conjugates of $h_{\mathbf{0}}$ under $GSp_{2n}(\mathbb{R})$, and write X^+ for the connected component of X containing $h_{\mathbf{0}}$. Then we have $X \cong \mathfrak{H}_n \sqcup \overline{\mathfrak{H}}_n$ by sending $g h_{\mathbf{0}} g^{-1}$ to $g(\sqrt{-1} \cdot 1_n) \in \mathfrak{H}_n$, where

$\mathfrak{H}_n = \{z \in M_n(\mathbb{C}) | {}^t z = z,\ \sqrt{-1}(z - \overline{z}) > 0\}$ (the Siegel upper half-space) and $\overline{\mathfrak{H}}_n$ (the Siegel lower half-space) is the complex conjugate of $\mathfrak{H}_n$. The action of $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in GSp_{2n}(\mathbb{R})$ (with $n \times n$ matrix a) on $\mathfrak{H}_n$ is given by $z \mapsto (az + b)(cz + d)^{-1}$.

For a scheme $S_{/R}$ defined over a finite $\mathcal{B}$-algebra R, when we regard it as a $\mathcal{B}$-scheme, we write it as $S_{/\mathcal{B}}$. The scheme $S_{/\mathcal{B}}$ is not equal to the Weil restriction of scalars $\mathrm{Res}_{R/\mathcal{B}} S_{/R}$ discussed in Chapter 4. Then we have $\mathfrak{M}^{J_n}_{N/\mathbb{Z}[\frac{1}{N}]}(\mathbb{C}) \cong G(\mathbb{Q})\backslash(X \times G(\mathbb{A}^{(\infty)}))/\widehat{\Gamma}(N)$, where we put

$$\widehat{\Gamma}(N) = \left\{\alpha \in GSp_{2n}(\widehat{\mathbb{Z}}) \middle| \alpha - 1 \in N \cdot M_{2n}(\widehat{\mathbb{Z}})\right\}.$$

We define, for $\mathbb{Z}[\zeta_N, \frac{1}{N}]$-scheme $\mathfrak{M}^{J_n}_{N/\mathbb{Z}[\zeta_N, \frac{1}{N}]}$,

$$Sh(G,X)_{/\mathbb{Q}} = \varprojlim_N \mathfrak{M}^{J_n}_{N/\mathbb{Z}[\frac{1}{N}]},\ Sh^{(p)}(G,X)_{/\mathbb{Z}_{(p)}} = \varprojlim_{p\nmid N} \mathfrak{M}^{J_n}_{N/\mathbb{Z}[\frac{1}{N}]}.$$

Therefore the construction of the Shimura variety for (G, X) follows from the existence theorem (Theorem 6.20) of the moduli variety over $\mathbb{Q}$ of polarized abelian varieties (proven by Mumford [GIT] and Shimura [Sh3]).

In this section, we give in this section an exposition on its global reciprocity law and the irreducibility of the Igusa tower over $Sh^{(p)}(G, X)$.

6.4.1 Moduli Functors

We consider the $\mathbb{Q}$-vector space $V = \mathbb{Q}^{2n}$ with alternating form $\langle x, y\rangle = {}^t x J_n y$. Thus $G(R) = GSp_{2n}(R)$ acts on $V_R = V \otimes_{\mathbb{Q}} R$ by matrix multiplication preserving the alternating form up to scalar multiplication. For each $(A, \lambda)_{/S} \in \mathcal{A}_{1,1}(S)$, choosing a geometric point $s = Spec(k(s))$ of each connected component of S, we consider its physical Tate module

$$\mathcal{T}_s(A) = \varprojlim_N A[N](k(s)),$$

which is naturally a module over $\widehat{\mathbb{Z}}$. If S is a $\mathbb{Q}$-scheme, $\mathcal{T}_s(A)$ is a free $\widehat{\mathbb{Z}}$-module of rank $2n$. If S is a $\mathbb{Z}_{(p)}$-scheme, its prime-to-p part $\mathcal{T}_s^{(p)}(A) = \mathcal{T}_s(A) \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}^{(p)}$ is a free $\widehat{\mathbb{Z}}^{(p)}$-module of rank $2n$. We then define

$$V_s(A) = \mathcal{T}_s(A) \otimes_{\mathbb{Z}} \mathbb{A}^{(\infty)} \quad \text{and} \quad V_s^{(p)}(A) = \mathcal{T}_s(A) \otimes_{\mathbb{Z}} \mathbb{A}^{(p\infty)}.$$

We consider level structures $\eta : V_{\mathbb{A}^{(\infty)}} \cong V_s(A)$ and $\eta^{(p)} : V_{\mathbb{A}^{(p\infty)}} \cong V_s^{(p)}(A)$. An element $g \in G(\mathbb{A})$ acts on η (and on $\eta^{(p)}$) by $\eta \mapsto \eta \circ g$. For a closed subgroup $K \subset G(\mathbb{A}^{(\infty)})$, the level K-structure $\overline{\eta} = \eta \circ K$ is defined over S, if $\sigma \circ \eta = \eta \circ k_s(\sigma)$ with $k_s(\sigma) \in K$ for all $\sigma \in \pi_1(S, s)$ for each chosen s. Changing s to another point s' in the same connected component gives isomorphisms (called paths) $\iota_{s,s'} : \pi_1(S, s) \cong \pi_1(S, s')$ and $\iota^{\mathcal{T}}_{s,s'} : \mathcal{T}_s(A) \cong \mathcal{T}_{s'}(A)$ which satisfy the compatibility condition (see Section 4.4 in the text

and [ECH] Section I.5) $\iota^{\mathcal{T}}_{s,s'}(\sigma(x)) = \iota_{s,s'}(\sigma)(\iota^{\mathcal{T}}_{s,s'}(x))$; so, if $\overline{\eta}$ is defined over S, we have

$$\iota^{\mathcal{T}}_{s,s'}(\eta \circ k_s(\sigma)v) = \iota^{\mathcal{T}}_{s,s'}(\sigma \circ \eta(v)) = \iota_{s,s'}(\sigma) \circ (\iota^{\mathcal{T}}_{s,s'}(\eta(v))).$$

Thus giving $\overline{\eta}$ defined over S at s is equivalent to giving $\iota^{\mathcal{T}}_{s,s'} \circ \overline{\eta}$ again defined over S at s'; so, we may forget about the choice s.

Since a positive multiple of the polarization is induced by an ample line bundle globally over S, the Galois action of $\pi_1(S,s)$ preserves the alternating form $\langle \cdot, \cdot \rangle_{\lambda,s}$ on $\mathcal{T}_s(A)$ induced by λ and the e_N-pairing on $A[N] \times {}^tA[N]$. The pairing $\langle \cdot, \cdot \rangle_{\lambda,s}$ is sent to $\langle \cdot, \cdot \rangle_{\lambda,s'}$ by $\iota^{\mathcal{T}}_{s,s'}$; so, the following condition

$$\langle \eta(x), \eta(y) \rangle_{\lambda,s} = \langle x, y \rangle \ \text{ up to the scalar multiple in } (\mathbb{A}^{(\infty)})^\times \tag{6.11}$$

is independent of the choice of s. Thus hereafter we drop the subscript s since it is irrelevant to our purpose. We always impose this condition on η in this section.

We consider the following functor $\mathcal{E}^{(\emptyset)}_K : \mathbb{Q}\text{-}SCH \to SETS$ for each closed subgroup K of $G(\mathbb{A}^{(\infty)})$,

$$\mathcal{E}^{(\emptyset)}_K(S) = \left[(A, \overline{\lambda}, \overline{\eta} = \eta \circ K)_{/S} \middle| (A,\lambda) \in \mathcal{A}_{1,1}(S),\ \eta : V_{\mathbb{A}^{(\infty)}} \cong V(A) \right], \tag{6.12}$$

where $\overline{\lambda}$ is a polarization class of polarizations proportional to a principal polarization on A up to multiplication by $\mathbb{Q}^\times_+$, and the alternating form induced by λ on $V_{\mathbb{A}^{(\infty)}}$ is proportional to $\langle \cdot, \cdot \rangle$ on $V_{\mathbb{A}^{(\infty)}}$. Here the isomorphism classes are taken in the category of abelian schemes up to isogenies; so, we have an isomorphism $(A, \overline{\lambda}, \overline{\eta}) \xrightarrow{\sim} (A', \overline{\lambda}', \overline{\eta}')$ if we have an isogeny $\phi : A \to A'$ with $\phi^* \overline{\lambda}' = \overline{\lambda}$ and $\overline{\eta}' = \phi \circ \overline{\eta}$ up to a nonzero scalar multiple. When $K = \widehat{\Gamma}(N)$, we simply write $\mathcal{E}^{(\emptyset)}_N$ for $\mathcal{E}^{(\emptyset)}_K$. In the definition of the functor $\mathcal{E}^{(\emptyset)}_K$, we do not need the condition similar to (pol) in Section 4.2.1 because we can choose an isomorphism $V \cong H_1(A, \mathbb{Q})$ for each test object $(A, \overline{\lambda}, \overline{\eta})$ compatible with $\eta : V_{\mathbb{A}^{(\infty)}} \cong V(A) = H_1(A, \mathbb{A}^{(\infty)})$ (because alternating forms isomorphic over $\mathbb{A}^{(\infty)}$ are isomorphic over $\mathbb{Q}$; see Section 7.1.5).

Let $L = \mathbb{Z}^{2n} \subset V$ be the self-dual lattice under $\langle \cdot, \cdot \rangle$. We write $\widehat{L}$ for $L \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}$. Then in the same manner as in Section 4.2.1, we can show an isomorphism of functors $\mathcal{E}^{(\emptyset)}_K \cong \mathcal{E}_K$ (for $K \subset G(\widehat{\mathbb{Z}})$), where

$$\mathcal{E}_K(S) = \left\{ (A, \lambda, \overline{\eta})_{/S} \middle| \eta(\widehat{L}) = \mathcal{T}(A) \right\} / \cong .$$

Here we have an isomorphism $(A, \lambda, \overline{\eta}) \overset{\phi}{\cong} (A', \lambda', \overline{\eta}')$ if $\phi : A \xrightarrow{\sim} A'$ is an isomorphism (not an isogeny) of abelian schemes with $\phi^* \lambda' = \lambda$ and $\overline{\eta}' = \phi \circ \overline{\eta}$. When $K = \widehat{\Gamma}(N)$, we again write $\mathcal{E}_N$ for $\mathcal{E}_K$. Then $\mathcal{E}_N$ is represented by $\mathfrak{M}^J_{N/\mathbb{Q}}$, because having $\eta \mod \widehat{\Gamma}(N)$ with $\eta(\widehat{\mathbb{Z}}^2) = \mathcal{T}(A)$ (so that the Weil pairing inducing the polarization form $\langle \cdot, \cdot \rangle$) is equivalent to having $\phi_N : (\mathbb{Z}/N\mathbb{Z})^{2n} =$

$L/NL \cong A[N]$ (compatible with the polarization form $\langle \cdot, \cdot \rangle$). Thus $\mathcal{E}_1^{(\emptyset)}$ is represented by $Sh(G,X)_{/\mathbb{Q}} = \varprojlim_N \mathfrak{M}^J_{N/\mathbb{Q}}$.

In order to construct a p-integral model of the Shimura variety, we impose (6.11) by substituting $\eta^{(p)}$ for η and require that $\overline{\lambda}$ be a polarization class of a principal polarization up to a prime-to-p scalar multiple. Then we can consider the following smooth model $Sh^{(p)}(G,X)_{/\mathbb{Z}_{(p)}} = \varprojlim_{p\nmid N} \mathfrak{M}^J_{N/\mathbb{Z}_{(p)}}$, whose quotient $Sh^{(p)}_K = Sh^{(p)}(G,X)/K$ by a closed subgroup K with $K = G(\mathbb{Z}_p) \times K^{(p)}$ ($K^{(p)} \subset G(\mathbb{A}^{(p)})$) represents the following functor $\mathcal{E}^{(p)}_K : \mathbb{Z}_{(p)}\text{-}SCH \to SETS$,

$$\mathcal{E}^{(p)}_K(S) = \left[(A, \overline{\lambda}, \overline{\eta}^{(p)})_{/S} \middle| \eta^{(p)} : V_{\mathbb{A}^{(p\infty)}} \cong V^{(p)}(A) \right], \tag{6.13}$$

where $\overline{\eta}$ indicates the $K^{(p)}$ orbit $\eta^{(p)} \circ K^{(p)}$. Here the isomorphism classes are taken in the category of abelian schemes up to prime-to-p isogeny; so, we have an isomorphism $(A, \overline{\lambda}, \overline{\eta}^{(p)}) \xrightarrow{\sim} (A', \overline{\lambda}', \overline{\eta}'^{(p)})$ if we have an isogeny $\phi : A \to A'$ with $p \nmid |\mathrm{Ker}(\phi)| = \deg(\phi)$, $\phi^* \overline{\lambda}' = \overline{\lambda}$, and $\overline{\eta}'^{(p)} = \phi \circ \overline{\eta}^{(p)}$ up to a prime-to-p scalar multiple.

6.4.2 Siegel Modular Reciprocity Law

We can let $g \in G(\mathbb{A}^{(\infty)})$ act on the functor $\mathcal{E}_1^{(\emptyset)}$ by $(A, \overline{\lambda}, \eta)_{/S} \mapsto (A, \overline{\lambda}, \eta \circ g)_{/S}$, and hence we get a schematic action of $G(\mathbb{A}^{(\infty)})$ on $Sh(G,X)$. Since the isogeny $\xi \in \mathrm{End}^{\mathbb{Q}}(A_{/S})$ for $\xi \in Z(\mathbb{Q})$ induces an isomorphism $(A, \overline{\lambda}, \eta) \cong (A, \overline{\lambda}, \eta \circ \xi)$ in the category of abelian schemes up to isogeny, $Z(\mathbb{Q})$ acts trivially on $Sh(G,X)$, and the action factors through $\frac{G(\mathbb{A}^{(\infty)})}{Z(\mathbb{Q})}$. Let Σ be a proper subset of rational primes. We define $\overline{\mathcal{E}}^{(\Sigma)}(G,X) = \frac{G(\mathbb{A}^{(\Sigma\infty)})}{Z(\mathbb{Z}_{(\Sigma)})}$. When $\Sigma = \emptyset$, we have

$$\overline{\mathcal{E}}^{(\emptyset)}(G,X) = \frac{G(\mathbb{A}^{(\infty)})}{Z(\mathbb{Q})} = \frac{G(\mathbb{A}^{(\infty)})}{\overline{Z(\mathbb{Q})}} *_{G(\mathbb{Q})_+} G^{ad}(\mathbb{Q})$$

in terms of Deligne's formulation described in Section 4.2.2. Note that $Z(\mathbb{Q}) = \mathbb{Q}^\times$ is closed in $G(\mathbb{A}^{(\infty)})$ (a property special to $\mathbb{Q}$); so, $\overline{Z(\mathbb{Q})} = Z(\mathbb{Q})$. We put $\mathfrak{M}^{(\Sigma)}_{/\mathbb{Q}^{cyc}} = \varprojlim_N \mathfrak{M}^J_{N/\mathbb{Q}[\mu_N]}$ for N running integers prime to Σ as a subvariety of $Sh^{(\Sigma)} = Sh(G,X)/G(\mathbb{Z}_\Sigma)$. Recall that $\mathbb{Q}^{(\Sigma)}/\mathbb{Q}$ is the maximal extension inside $\mathbb{Q}^{cyc}$ unramified at every $p \in \Sigma$. Here is the global reciprocity law.

Theorem 6.26 (Shimura) *Let $G = GSp(2n)_{/\mathbb{Q}}$. Then the variety $\mathfrak{M}^{(\Sigma)}$ over $\mathbb{Q}^{(\Sigma)}$ gives a geometrically irreducible component of $Sh^{(\Sigma)}(G,X)_{/\mathbb{Q}}$ containing the image of $\mathfrak{H}_n \times 1$. Writing $\mathfrak{K}^{(\Sigma)}$ for the function field $\mathbb{Q}^{(\Sigma)}(\mathfrak{M}^{(\Sigma)})$, we have the identity $\mathrm{Aut}(\mathfrak{K}^{(\Sigma)}) = \mathrm{Aut}(\mathfrak{M}^{(\Sigma)}_{/\mathbb{Q}})$, and we have a canonical isomorphism $\tau : \overline{\mathcal{E}}^{(\Sigma)}(G,X) \xrightarrow{\sim} \mathrm{Aut}(\mathfrak{K}^{(\Sigma)})$ satisfying the following properties.*

(1) *On* $\mathbb{Q}^{(\Sigma)}$, $g \in G(\mathbb{A}^{(\Sigma\infty)})$ *acts by the Artin reciprocity map in* Theorem 2.12 *composed with the similitude norm* $\nu : G(\mathbb{A}^{(\Sigma\infty)}) \to \mathbb{A}^\times$;
(2) *For* $f \in \mathfrak{K}^{(\Sigma)}$ *regarded as a function on* X^+, *we have* $\tau(\alpha)(f) = f \circ \alpha^{-1}$ *for* $\alpha \in G(\mathbb{Z}_{(\Sigma)})_+$;
(3) *The map* τ *is an isomorphism of topological groups with respect to the adele topology on* $\overline{\mathcal{E}}^{(\Sigma)}(G, X)$ *and the Krull topology on* $\mathrm{Aut}(\mathfrak{K}^{(\Sigma)})$.

The action of $\tau(\alpha)$ looks different from the one in Theorem 2.43 and actually, our $\tau(x)$ here is the inverse of $\tau(x)$ in Theorem 2.43 in order to make the construction of τ compatible with our definition of the action on the level structure $\eta \mapsto \eta \circ g$ (see (6.14) below).

Recall the valuation ring $\mathbb{Z}_{(p)}^{(\Sigma)}$ (with residue field $\mathbb{F}$) in $\mathbb{Q}^{(\Sigma)}$ of the p-adic place $i_p : \mathbb{Q}^{(\Sigma)} \hookrightarrow \overline{\mathbb{Q}}_p$. If $p \in \Sigma$, we have a modulo p version of this theorem determining $\mathrm{Aut}(\mathfrak{k}_p^{(\Sigma)})$ for the function field $\mathfrak{k}_p^{(\Sigma)}$ of $\mathfrak{M}_{/\mathbb{F}}^{(\Sigma)} = \mathfrak{M}^{(\Sigma)} \otimes_{\mathbb{Z}_{(p)}^{(\Sigma)}} \mathbb{F}$. Since it is similar to Theorem 4.17, we leave the formulation of the result to the reader (see [H03b]).

We repeat here the argument given for the Hilbert modular version (Theorem 4.14) based on an argument in [AAF] Section 8 and the theorem of Faltings and Zarhin (Theorem 4.13). For the original proof of Shimura (and Miyake), we refer the reader to [Sh5] II 6.5 and [Mt].

Proof. Since the proof is the same for any choice of Σ, we assume that $\Sigma = \emptyset$ for simplicity, and we throw all superscripts and subscripts Σ away from our notation. By our convention, $\mathbb{Z}_{(\emptyset)} = \mathbb{Q}$. We first check that $\tau(\alpha^{-1})(f) = f \circ \alpha$ by studying the complex points $\mathfrak{M}_N^J(\mathbb{C})$ of $\mathfrak{M}_N^J$.

For a given $(A, \lambda) \in \mathcal{E}_1(\mathbb{C})$, we choose a base $\{\omega_1, \ldots, \omega_n\}$ of $H^0(A, \Omega_{A/\mathbb{C}})$, and consider the lattice $\mathcal{L}_A = \{(\int_\gamma \omega_j)_j | \gamma \in H_0(A, \mathbb{Z})\}$ in $\mathbb{C}^n$. By the Poincaré duality and the principal polarization λ, we may assume that

$$\mathcal{L}_A = \mathcal{L}_z = {}^tL \cdot J_n \begin{pmatrix} z \\ 1_n \end{pmatrix}$$

for $z \in \mathfrak{H}_n$ (see [AAF] 4.4), regarding L to be a column vector space. Then $A(\mathbb{C}) \cong \mathbb{C}^n/\mathcal{L}_z$ as complex manifolds, and λ is induced by the alternating form $\lambda_z : \mathcal{L}_z \wedge \mathcal{L}_z \to \mathbb{Z}$ sending ${}^tx \cdot J_n \begin{pmatrix} z \\ 1_n \end{pmatrix} \wedge {}^ty \cdot J_n \begin{pmatrix} z \\ 1_n \end{pmatrix}$ to $x \cdot J_n {}^ty \in \mathbb{Z}$. We write A_z for the abelian variety with $A_z(\mathbb{C}) \cong \mathbb{C}^n/\mathcal{L}_z$. Define the main involution ι of G by $\alpha^\iota = J_n^{-1} \cdot {}^t\alpha J_n$ for $\alpha \in G(R)$. Note that $\langle \alpha x, y \rangle = \langle x, \alpha^\iota y \rangle$ for the alternating form $\langle x, y \rangle = {}^tx \cdot J_n y$ on L.

If we have an isogeny $\phi : A(\mathbb{C}) \cong \mathbb{C}^n/\mathcal{L}_z \to \mathbb{C}^n/\mathcal{L}_w \cong Y(\mathbb{C})$ preserving the polarization class, and identifying $\mathbb{C}^n$ with the Lie algebra of the abelian varieties, we have a $\mathbb{C}$-linear map $\Phi : \mathbb{C}^n \to \mathbb{C}^n$ lifting ϕ with $\Phi(\mathbb{Q}\mathcal{L}_z) = \mathbb{Q}\mathcal{L}_w$ preserving the polarization form. In other words, we find $\alpha \in GL_{2n}(\mathbb{Q})$ such that

$$\Phi(\ell \cdot J_n \begin{pmatrix} z \\ 1_n \end{pmatrix}) = \ell \cdot \alpha \cdot J_n \begin{pmatrix} w \\ 1_n \end{pmatrix} j(\alpha^\iota, z)$$

for all $\ell \in L$, where $\alpha^\iota = J_n^{-1} \cdot {}^t\alpha J_n = \nu(\alpha)\alpha^{-1}$ and $j(g, z) = cz + d \in GL_n(\mathbb{C})$ for $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Since α preserves polarization up to positive rational numbers, $\nu(\alpha) > 0$, and hence $\alpha \in G(\mathbb{Q})_+$. This shows that $w = \alpha(z)$ for $\alpha \in G(\mathbb{Q})_+$.

We define $\phi_{N,z} : L/NL \cong A_z[N]$ by $\phi_{N,z}(x) = \frac{{}^t x}{N} J_n \cdot \left(\begin{smallmatrix} z \\ 1_n \end{smallmatrix}\right) \mod \mathcal{L}_z$. Then $(A_z, \lambda, \phi_{N,z}) \cong (A_w, \lambda, \phi_{N,w}) \Leftrightarrow w = \gamma(z)$ for some $\gamma \in \Gamma(N) = \widehat{\Gamma}(N) \cap G(\mathbb{Q})_+$. For any given level structure $\phi : L/NL \cong A_z[N]$ compatible with the polarization λ, we find $\gamma \in Sp_{2n}(\mathbb{Z})$ such that $\phi = \phi_{N,z} \circ \gamma$. Thus we find $\mathfrak{M}_N^J(\mathbb{C}) \cong \Gamma(N)\backslash\mathfrak{H}_n$, in particular, $\mathfrak{M}_N^J$ over $\mathbb{Z}[\zeta_N, \frac{1}{N}]$, is geometrically irreducible and smooth if $N \geq 3$. This follows from the same argument in the Hilbert modular case (given just above Theorem 4.17), using the existence of the smooth projective toroidal compactification of Faltings–Chai described in Section 6.3.6, by Zariski's connectedness theorem ([ALG] III.11.3). This shows that $\mathfrak{M}_{/\mathbb{Q}^{cyc}}$ gives the geometrically irreducible component of $Sh(G,X)_{/\mathbb{Q}^{cyc}}$ containing $X^+ \times 1 = \mathfrak{H}_n \times 1$.

By taking the limit $\phi_z = \varprojlim_N \phi_{N,z} : \widehat{L} \cong \mathcal{T}(A_z)$ and tensoring $\mathbb{A}^{(\infty)}$, we get $\eta_z : V_{\mathbb{A}^{(\infty)}} \cong V(A_z)$. We let $\alpha \in G(\mathbb{Q})_+$ act on column vectors in $V = L \otimes_{\mathbb{Z}} \mathbb{Q}$ by (left) multiplication. Since the linear map $j(\alpha^\iota, z) : \mathbb{C}^n \ni u \mapsto j(\alpha^\iota, z)u \in \mathbb{C}^n$ for $\alpha \in G(\mathbb{Q})_+$ induces an isogeny

$$(A_{\alpha^{-1}(z)}, \lambda_{\alpha^{-1}(z)}, \phi_{N,\alpha^{-1}(z)}) \to (\mathbb{C}^n/{}^t L^t \alpha J_n \left(\begin{smallmatrix} z \\ 1_n \end{smallmatrix}\right), j(\alpha^\iota, z)_* \lambda_z, j(\alpha^\iota, z)\eta_{\alpha^{-1}(z)}) \cong (A_z, \lambda_z, \eta_z \circ \alpha), \quad (6.14)$$

we have $\tau(\alpha^{(\infty)})([z,1]) = [z, \alpha^{(\infty)}] = [\alpha^{-1}(z), 1]$, writing $[z,g]$ for the complex point of $Sh(G,X)(\mathbb{C})$ represented by $(z \times g) \in \mathfrak{H}_n \times G(\mathbb{A}^{(\infty)})$ as in (4.21). This shows that $\alpha \in G(\mathbb{Q})_+ \subset G(\mathbb{A}^{(\infty)})$ acts on $f \in \mathfrak{K}$ by $f \mapsto f \circ \alpha^{-1}$.

By the above expression of the complex point of $\mathfrak{M}$, we find that τ brings $G_1(\mathbb{A}^{(\infty)})/\{\pm 1\}$ injectively into $\mathrm{Aut}(\mathfrak{M}_{/\mathbb{Q}})$. We consider the $e_{N,z}$-pairing $e_{N,z} : A_z[N] \times A_z[N] \to \mu_N$. Since $e_{N,z}(\phi_{N,z}(x), \phi_{N,z}(y)) = \zeta_N^{\langle x,y\rangle}$ for $x, y \in L/NL$, we find that $e_{N,z}(\alpha x, \alpha y) = \zeta_N^{\nu(\alpha)\langle x,y\rangle}$ for $\alpha \in G(\widehat{\mathbb{Z}})$. Then in the same manner as in the proof of Theorem 4.14 Step 2, we create a section $s_{\mathbb{A}} : \widehat{\mathbb{Z}}^\times \ni c \mapsto \left(\begin{smallmatrix} c1_n & 0 \\ 0 & 1_n \end{smallmatrix}\right) \in G(\mathbb{A}^{(\infty)})$ of ν so that the action of $\tau(s_{\mathbb{A}}(c))$ on the q-expansion at ∞ factors through the Galois action of the Artin symbol $[c, \mathbb{Q}]$ on the coefficients of the q-expansion. Thus we get the following commutative diagram with exact rows (and the section $s_{\mathbb{A}}$ of ν),

$$\begin{array}{ccccc} \mathrm{Aut}(\mathfrak{M}_{/\mathbb{Q}^{cyc}}) & \hookrightarrow & \mathrm{Aut}(\mathfrak{M}_{/\mathbb{Q}}) & \twoheadrightarrow & \mathrm{Gal}(\mathbb{Q}^{cyc}/\mathbb{Q}) \\ \cup\uparrow & & \uparrow\tau & & \wr\uparrow\iota \\ G_1(\mathbb{A}^{(\infty)})/\pm 1 & \hookrightarrow & \frac{G(\mathbb{A}^{(\infty)})}{Z(\mathbb{Q})} & \xrightarrow{\nu} & \mathbb{A}^\times/\mathbb{Q}^\times\mathbb{R}_+^\times = \widehat{\mathbb{Z}}^\times, \end{array}$$

where ι is the Artin reciprocity map in Theorem 2.12.

After verifying this, we proceed in exactly the same manner as in Step 3 of the proof of Theorem 4.14, and we get surjectivity of τ out of Theorem 4.13 combined with the results shown in Section 4.4.3. □

The local reciprocity law at CM points is similar to Corollary 4.20 (see (rc1–2) in Section 7.2.2 and Theorem 7.7).

6.4.3 Siegel Modular Igusa Tower

We can make the toroidal compactification M_N of $\mathfrak{M}_N^J$ (for N prime to p) choosing the toroidal compactification data carefully so that the group $G(\widehat{\mathbb{Z}}^{(p)})$ acts on the limit $M^{(p)}_{/\mathbb{Z}_{(p)}} = \varprojlim_{p\nmid N} M_{N/\mathbb{Z}_{(p)}}$. This is done as follows. Choose small $N_0 \geq 3$ and a sequence $N_0|N_1|N_2|\cdots|N_j|N_{j+1}|\cdots$ of positive integers so that $Sh^{(p)}(G,X) = \varprojlim_j \mathfrak{M}_{N_j}$. Fix a simplicial decomposition of C for N_0 at the cusp ∞. We take C sufficiently fine so that M_{N_0} is smooth projective over $\mathbb{Z}[\zeta_{N_0}, \frac{1}{N_0}]$. Then we take the isomorphic image of the decomposition at every other cusp. Write this decomposition as C_s for each cusp $s \in M^*_{N_0}$. Using the same decomposition C_s for all cusps of M_{N_j} over s, we still get a smooth projective compactification M_{N_j} which is a finite Galois covering with Galois group $\widehat{\Gamma}(N_0)/\widehat{\Gamma}(N_j)$ as described in Theorem 6.24. Since the action of $Sp_{2n}(\mathbb{Z}/N_j\mathbb{Z})$ over $\mathfrak{M}^J_{N_j}$ permutes cusps, by our choice, the geometric action of $\widehat{\Gamma}(N_0)/\widehat{\Gamma}(N_j)$ on M_{N_j} extends to the action of $Sp_{2n}(\mathbb{Z}/N_j\mathbb{Z})$ on M_{N_j}. The arithmetic action of $GSp_{2n}(\mathbb{Z}/N_j\mathbb{Z})/Sp_{2n}(\mathbb{Z}/N_j\mathbb{Z}) \cong \mathrm{Gal}(\mathbb{Q}(\zeta_{N_j})/\mathbb{Q})$ is through coefficients $\mathbb{Z}[\zeta_N, \frac{1}{N}]$, which extends to $M_{N_j/\mathbb{Z}_{(p)}}$ as described in the proof of the global reciprocity law (Theorem 6.26). Thus $GSp_{2n}(\mathbb{Z}/N_j\mathbb{Z})$ acts on $M_{N_j/\mathbb{Z}_{(p)}}$. The action of $GSp_{2n}(\mathbb{Z}/N_{j+1}\mathbb{Z})$ is compatible with the projection map $M_{N_{j+1}} \twoheadrightarrow M_{N_j}$, because the polyhedral decomposition for the level N_{j+1} is the same as that for the level N_j. Taking the limit, we get a good action of $G(\widehat{\mathbb{Z}}^{(p)})$ on $M^{(p)}_{/\mathbb{Z}_{(p)}}$.

The limit $M^{(p)} = \varprojlim_j M_{N_j/\mathbb{Z}_{(p)}}$ gives a smooth compactification over $\mathbb{Z}_{(p)}$ of $Sh^{(p)}$ keeping the action of $G(\widehat{\mathbb{Z}}^{(p)})$. We then define $M^{(p)}_{K/\mathbb{Z}_{(p)}} = M^{(p)}/K$ for a closed subgroup $K^{(p)} \subset G(\widehat{\mathbb{Z}}^{(p)})$ with $K = G(\mathbb{Z}_p) \times K^{(p)}$ (i.e., a closed subgroup maximal at p). In particular, $M^{(p)}/\widehat{\Gamma}(N_j) = M_{N_j/\mathbb{Z}_{(p)}}$ for all j. By minimality, on the limit ${M^{(p)}}^* = \varprojlim_{p\nmid N} M^*_{N/\mathbb{Z}_{(p)}}$ of the minimal compactifications, the adele group $G(\mathbb{A}^{(p\infty)})$ acts naturally. In particular, ${M^{(p)}}^*_K = {M^{(p)}}^*/K$ is the minimal compactification of $Sh^{(p)}_K$.

We go through the same process for $Sh(G,X)_{/\mathbb{Q}} = \varprojlim_N \mathfrak{M}^J_{N/\mathbb{Q}}$ and get a toroidal compactification $M_{/\mathbb{Q}} = \varprojlim_N M_{N/\mathbb{Q}}$ which has the natural action of $G(\widehat{\mathbb{Z}})$ extending that on $Sh(G,X)$. We also define $M^*_{/\mathbb{Q}} = \varprojlim_N M^*_{N/\mathbb{Q}}$ on which $G(\mathbb{A}^{(\infty)})$ acts. In this case, we have a smooth toroidal compactification $M_{K/\mathbb{Q}} = M/K$ and the minimal compactification $M^*_{K/\mathbb{Q}} = M^*/K$ for all closed subgroups K of $G(\mathbb{A}^{(\infty)})$.

We have defined in Section 4.1.6 the Hasse invariant H for any semi-abelian scheme $f : \mathcal{G} \to S$ over a characteristic p-scheme with a given base of invariant differentials and verified that $H \in \det \underline{\omega}^{\otimes p-1}$ where $\det \underline{\omega} = f_*(\bigwedge^n \Omega_{\mathcal{G}/S})$. In particular, H vanishes at any point of S where $\mathcal{G}$ has a nonordinary semi-abelian scheme as its fiber. Here a semi-abelian scheme is p-ordinary if the connected component $\mathcal{G}[p]^\circ$ of the kernel $\mathcal{G}[p]$ of multiplication by p is iso-

morphic to μ_p^n étale locally. If $\mathcal{G}$ is an abelian scheme, by the duality, this is equivalent to having the étale quotient $\mathcal{G}[p]^{et}$ isomorphic to $(\mathbb{Z}/p\mathbb{Z})^n$ étale locally. Thus we have a well-defined Hasse invariant $H \in \det(\underline{\omega})^{\otimes p-1}_{/M^{(p)}}$ because $M^{(p)}$ carries a semi-abelian scheme (depending on the toroidal compactification data).

We pick a lift $E \in H^0(M^{(p)}_{/\mathbb{Z}_{(p)}}, \det(\underline{\omega})^{a(p-1)})$ (for a suitable $0 < a \in \mathbb{Z}$) of the Hasse invariant H of the (universal) semi-abelian scheme $\mathcal{G}$ over the toroidal compactification $M^{(p)}_{/\mathbb{F}_p}$. Then we put $S_m = (M^{(p)} \otimes_{\mathbb{Z}_{(p)}} \mathbb{Z}/p^m\mathbb{Z})[\frac{1}{E}]$, $S^*_m = (M^{*(p)} \otimes_{\mathbb{Z}_{(p)}} \mathbb{Z}/p^m\mathbb{Z})[\frac{1}{E}]$, $S_{K,1} = S_1/K$, and $S^*_{K,1} = S^*_1/K$ all defined over $\mathbb{F}_p$. Since $M^*_K \otimes \mathbb{F} = Proj(\mathfrak{G}_K)$ for $\mathfrak{G}_K = \bigoplus_{0 \le_J} H^0(Sh^{(p)}_{K/\mathbb{F}_p}, \det \underline{\omega}^{\otimes_J}_{/\mathbb{F}_p})$, we have $S^*_{K,1} = Proj(\mathfrak{G}_K[\frac{1}{H}]) = Spec(\mathfrak{G}_K[\frac{1}{H}]_0)$ (cf. [GME] 1.3.2), which is affine. Although $M_{K/\mathbb{F}_p}$ is projective, $\det \underline{\omega}$ is not ample on M_K; so, $S_{K,1}$ is not necessarily affine.

We define the self-product $\mathcal{G}^n_K = \overbrace{\mathcal{G}_K \times_{M_K} \mathcal{G}_K \times \cdots \times_{M_K} \mathcal{G}_K}^{n}$. Here $\mathcal{G}_K$ is the semi-abelian scheme over M_K, which is again given by $\mathcal{G}/K$ for the limit $\mathcal{G}_{/M^{(p)}}$ of semi-abelian scheme $\mathcal{G}_{N_j}$ over M_{N_j}. Then the Cartier dual A_m of the connected component $\mathcal{G}^n_K[p^m]^\circ$ of $\mathcal{G}^n_K[p^m]$ over $S_{K,1}$ is an additive $M_n(\mathbb{Z}/p^m\mathbb{Z})$-torsor, because we have $\mathcal{G}^n_K[p^m]^\circ \cong \mu^{n^2}_{p^m}$ (étale locally) whose Cartier dual is $M_n(\mathbb{Z}/p^m\mathbb{Z})$. We take its multiplicative $GL_n(\mathbb{Z}/p^m\mathbb{Z})$-subtorsor $T_{K,m} = \{(g_1, \ldots, g_n) \in A_m | g_1 \wedge g_2 \wedge \cdots \wedge g_n \neq 0\}$. As we show later, this $T_{K,m}$ is the m-th layer of the Igusa tower over $S_{K,1/\mathbb{F}_p}$. Thus $T_{K,m}/S_{K,1}$ is an étale covering on which $GL_n(\mathbb{Z}/p^m\mathbb{Z})$ naturally acts. We prove the geometric irreducibility of $T_{K,m}$ over each connected component of $S_{K,1/\overline{\mathbb{F}}_p}$ for all open compact subgroups K maximal at p.

Let O be a ring. For two given group schemes (on which O acts linearly) $f : P \to S$ and $f' : P' \to S$, we define the following functors $\underline{\mathrm{Isom}}_S(P, P') : SCH_{/S} \to SETS$ and $\underline{\mathrm{Hom}}_S(P, P') : SCH_{/S} \to SETS$,

$$\begin{aligned} \underline{\mathrm{Hom}}_S(P, P')(S') &= \left\{\phi \in \mathrm{Hom}_{S'\text{-}SCH}(P_{S'}, P'_{S'}) \middle| \phi \text{ is } O\text{-linear}\right\} \\ \underline{\mathrm{Isom}}_S(P, P')(S') &= \left\{\phi \in \underline{\mathrm{Hom}}_S(P, P')(S') \;\middle|\; \phi \text{ is an isomorphism}\right\}, \end{aligned} \tag{6.15}$$

where $P_{S'} = P \times_S S'$. Let P_m be the Cartier dual of $\mathcal{G}_K[p^m]^\circ$ (see Section 3.1.6 for Cartier dual). Having a homomorphism $\phi : (\mathbb{Z}/p^m\mathbb{Z})^n_{/S'} \to P_{m/S'}$ is just equivalent to knowing the destination of the standard base $\{e_1, \ldots, e_n\}$ in P_m. By assigning $(\phi(e_1), \ldots, \phi(e_n)) \in A_m(S')$ to $\phi \in \mathrm{Hom}_S((\mathbb{Z}/p^m\mathbb{Z})^n, P_m)(S')$, the $M_n(\mathbb{Z}/p^m\mathbb{Z})$-torsor A_m represents the functor $\underline{\mathrm{Hom}}_S((\mathbb{Z}/p^m\mathbb{Z})^n, P_m)$. In the same way, $T_{K,m}$ represents $\underline{\mathrm{Isom}}_S((\mathbb{Z}/p^m\mathbb{Z})^n, P_m)$. By taking the Cartier dual and taking the inverse of elements in $\mathrm{Isom}_S(\mathcal{G}[p^m]^\circ, \mu^n_{p^m})$, $T_{K,m}$ also represents the functor $\underline{\mathrm{Isom}}_S(\mu^n_{p^m}, \mathcal{G}[p^m]^\circ)$. Each element $g \in GL_n(\mathbb{Z}/p^m\mathbb{Z})$ acts on $\phi \in \mathrm{Isom}_S((\mathbb{Z}/p^m\mathbb{Z})^n, P_m)(S')$ by $\phi \mapsto \phi \circ g$.

Let W be a complete discrete valuation ring of the p-adic completion of the maximal unramified extension of $\mathbb{Q}_p$ in $\overline{\mathbb{Q}}_p$. Put $\mathcal{W} = i_p^{-1}(W) \subset \overline{\mathbb{Q}}$. Thus the residue field $\mathbb{F}$ of W and $\mathcal{W}$ is an algebraic closure of $\mathbb{F}_p$.

Theorem 6.27 (Chai–Faltings) *Let K be an open compact subgroup of $G(\mathbb{A}^{(\infty)})$ maximal at p. Then the projection $\pi_0(T_{K,1}) \to \pi_0(S_{K,1})$ is bijective; so, the pullback in $T_{K,1}$ of an irreducible component of $S_{K,1}$ is irreducible. If $\nu(K) = \widehat{\mathbb{Z}}^\times$, then $S_{K,1/\mathbb{F}_p}$ and $T_{K,m/\mathbb{F}_p}$ are geometrically irreducible.*

This is, of course, a generalization of a result of Igusa (Theorem 3.3). This has been proven in [DAV] V.7 showing that the p-adic monodromy group at the cusp ∞ is large. We give a proof via the global reciprocity law. Our proof is based on the existence of the smooth toroidal compactification of the Siegel modular variety given by [DAV], but our viewpoint is slightly different from the treatment in [DAV] V.7, looking into the monodromy of the generic point. This viewpoint can be generalized to prove irreducibility of the (generalized) Igusa tower for general unitary and symplectic groups over totally real fields bigger than $\mathbb{Q}$ (see Section 8.4 and [H03a] Lecture 10), which cannot be attained by the computation in [DAV] of cuspidal monodromy (see the remark after Theorem 4.21).

Proof. We need to prove that $\mathcal{T}_{K,m} = T_{K,m} \times_{S_{K,1}} S_K$ is geometrically irreducible for each geometrically irreducible component $S_K \subset S_{K,1}$.

The valuation ring $\mathcal{W}$ is the strict Henselization of $\mathbb{Z}_{(p)}$ in $\overline{\mathbb{Q}}$ associated with $i_p : \overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_p$. Since all the assertions are geometric, we may extend the scalar of all the schemes involved from $\mathbb{Z}_{(p)}$ to $\mathcal{W}$ whose residue field is $\overline{\mathbb{F}}_p$, which we have written as $\mathbb{F}$. Since K is maximal at p, it contains $\widehat{\Gamma}(N)$ for some $N \geq 3$ prime to p, and we may assume that $K = \widehat{\Gamma}(N)$. Then M_K is smooth projective over $\mathcal{W}$, and hence by Zariski's connectedness theorem ([ALG] III.11.4), geometrically irreducible components of a special fiber are in bijection with those of the generic fiber (see the argument just above Theorem 4.17 applied to the Hilbert modular varieties). Thus a geometrically irreducible component of $\mathfrak{M}_{N,1} = \mathfrak{M}_N \otimes_{\mathcal{W}} \mathbb{F}$ is given by $\mathfrak{M}^J_N$ $(p \nmid N)$. Since $G(\mathbb{Z}/N\mathbb{Z})$ acts transitively on all other irreducible components, we may assume $S_N := S_{\widehat{\Gamma}(N)} = M^J_{N,1}[\frac{1}{E}]$. Write $S_{/\mathbb{F}} = \varprojlim_{p \nmid N} S_N$ which remains geometrically irreducible (because $\mathcal{W}$ contains all $\mathbb{Z}[\mu_N]$ for N prime to p). Let $S^\circ \subset S_{/\mathbb{F}}$ be the subscheme obtained from S removing the cuspidal divisors and $S^*_{/\mathbb{F}} = M^*[\frac{1}{E}]$ for the minimal compactification of $M^*_{/\mathbb{F}}$ of $\mathfrak{M}^J_{/\mathbb{F}}$.

We write $\mathcal{T}^\circ_{K,\alpha}$ for the scheme $\mathcal{T}_{K,\alpha} \times_{S_K} S^\circ_K$. Then write $\mathcal{T}^*_{K,\alpha}$ for the normalization of $S^*_{K/\mathbb{F}}$ in $\mathcal{T}_{K,\alpha/\mathbb{F}}$. Since S^*_K is normal at cusps, $\mathcal{T}^*_{K,\alpha}$ contains a $GL_n(\mathbb{Z}_p)$-orbit of infinity cusps. We choose a coherent system of the infinity cusps of $\mathcal{T}^*_{K,\alpha}$ (indexed by α) and write it as ∞. We have a tower of irreducible subvarieties $\{C_\alpha \subset \mathcal{T}_\alpha\}_\alpha$ such that the image of C_α in $\mathcal{T}^*_{K,\alpha}$ contains the infinity cusp. We may identify $\mathrm{Gal}(C_\infty/C_0)$ with a closed subgroup $H \subset GL_n(\mathbb{Z}_p)$. Thus we need to show that $H = GL_n(\mathbb{Z}_p)$.

We write the ring $\mathcal{B}[[S^2(X)]]$ for $X = \mathbb{Z}^n$ studied in Section 6.3.6 as a q-expansion ring $\mathcal{B}[[q^\xi]]_\xi$ for symmetric $n \times n$ matrices ξ. Here $\mathcal{B}$ is a base ring. Since we have a (rational) toroidal compactification of $\mathfrak{M}^J_{/\mathbb{Q}[\zeta_N]}$, for all $f \in H^0(\mathfrak{M}^J_{N/\mathcal{B}}, \det \underline{\omega}^{\otimes k})$ (with a $\mathbb{Q}[\zeta_N]$-algebra $\mathcal{B}$), we have its q-expansion

$f(q) = \sum_{\xi \geq 0} a(\xi, f) q^{\xi}$ with $a(\xi, f) \in \mathcal{B}$, where ξ runs over positive semi-definite $n \times n$ matrices with coefficients in $\mathbb{Q}$. The minimal compactification of $\mathfrak{M}^J_{N/\mathbb{Q}}$ is given by $Proj(\mathfrak{G}_{\widehat{\Gamma}(N)})$. Thus any element f in the function field $\mathfrak{K} = \mathbb{Q}^{cyc}(\mathfrak{M})$ for $\mathfrak{M} = \varprojlim_N \mathfrak{M}^J_{N/\mathbb{Q}[\mu_N]}$ can be written as $\frac{n}{d}$ with $n, d \in \mathfrak{G}_k$ for a suitable integer $k > 0$, where $\mathfrak{G}_k$ is the kth graded component of $\mathfrak{G}$. The minimal compactification is projective and can be embedded into a projective space using theta series, which has integral q-expansion coefficients (cf. [CSM] Appendix II and [AAF] Section 6). Thus for a given $f \in \mathfrak{K}$ or more generally f in the field of fractions $Frac(\mathfrak{G})$ of $\mathfrak{G}$, we can take n and d with q-expansion coefficients bounded under the p-adic valuation ord_p of $\mathcal{W}$. Thus, the valuation $v(n) = \mathrm{Inf}_{\xi}\, ord_p(a(\xi, n))$ is a well-defined element in $\mathbb{Q} \sqcup \{\infty\}$. Similarly $v(d) \in \mathbb{Q}$ is also well-defined. Then we define the valuation $v : Frac(\mathfrak{G}) \to \mathbb{Q} \sqcup \{\infty\}$ by $v(f) = v(n) - v(d)$.

We have an isomorphism $\tau : \overline{\mathcal{E}}(G, X) \xrightarrow{\sim} \mathrm{Aut}(\mathfrak{K}) = \mathrm{Aut}(\mathfrak{M}_{/\mathbb{Q}})$ of topological groups (Theorem 6.26). Since the restriction of $\tau(x)$ to $\mathbb{Q}^{cyc}$ is given by the action of the similitude norm $\nu(x) \in \mathbb{A}^{\times}/\mathbb{Q}^{\times}\mathbb{R}^{\times}_{+}$ composed with the Artin reciprocity (Theorem 2.12), τ induces another isomorphism $G_1(\mathbb{A}^{(\infty)})/\{\pm 1\} \xrightarrow{\sim} \mathrm{Aut}(\mathfrak{M}_{/\mathbb{Q}^{cyc}}) = \mathrm{Aut}(\mathfrak{K}/\mathbb{Q}^{cyc})$.

We write $\mathcal{D}_p$ for the decomposition group in $\mathrm{Aut}(\mathfrak{M}_{/\mathbb{Q}})$ for this valuation v. As seen in the proof of Theorem 3.3, $\mathcal{D}_p$ is a closed subgroup of $\tau(\overline{\mathcal{E}}(G, X))$, and we prove that $\mathcal{D}_p/I_p$ contains $GL_n(\mathbb{Z}_p)$ for the inertia subgroup I_p.

The scheme S^* is geometrically irreducible, that is, $S^* \otimes_{\mathcal{W}} \overline{\mathbb{Q}}$ and $S^* \otimes_{\mathcal{W}} \overline{\mathbb{F}}_p$ are both irreducible. We then construct $\mathcal{T}^{\circ}_{\alpha} = \varprojlim_K \mathcal{T}^{\circ}_{K,\alpha}$ and $\mathcal{T}^{*}_{\alpha} = \varprojlim_K \mathcal{T}^{*}_{K,\alpha}$, where $K \subset G(\mathbb{A}^{(\infty)})$ runs over all open compact subgroups maximal at p. Then $\mathcal{T}^{\circ}_{\infty}/S$ is an étale covering, and we take the geometrically irreducible component C^*_{∞} of $\mathcal{T}^*_{\infty}$ containing the infinity cusp, and put $C^{\circ}_{\infty} = C^*_{\infty} \cap \mathcal{T}^{\circ}_{\infty}$. Since the action of $GL_n(\mathbb{Z}_p)$ for the diagonal subgroup $GL(n)$ of G_1 commutes with the action of $G_1(\mathbb{A}^{(p\infty)})$ and coincides with the action of $GL_n(\mathbb{Z}_p)$ on $T^{\circ}_{\infty}/S^{\circ}$, we have a commutative diagram with exact rows,

$$\begin{array}{ccccc} \mathrm{Gal}(C^{\circ}_{\infty}/S^{\circ}) & \hookrightarrow & \mathrm{Aut}(C^{\circ}_{\infty}; S^{\circ}) & \xrightarrow{\mathrm{Res}} & \mathrm{Aut}(S^{\circ}) \\ \wr\downarrow & & \uparrow\cup & & \uparrow\cup \\ H & \hookrightarrow & (G_1(\mathbb{A}^{(p\infty)}))/\{\pm 1\} \times H & \twoheadrightarrow & G_1(\mathbb{A}^{(p\infty)})/\{\pm 1\}, \end{array}$$

where $\mathrm{Aut}(C^{\circ}_{\infty}; S^{\circ})$ is the group of automorphisms of C°_{∞} that induce automorphisms of S°. This shows that the irreducible component C°_{∞} is stable under the action of $g \in G_1(\mathbb{A}^{(p\infty)})$. In particular, the infinity cusp of C°_{∞} is sent by g to another cusp of the same irreducible component C^*_{∞}. This shows that $\tau(g)$ for $g \in G_1(\mathbb{A}^{(p\infty)})/\{\pm 1\} \subset \overline{\mathcal{E}}(G, X)$ preserves v.

There is another way to show that $G_1(\mathbb{A}^{(p\infty)})/\{\pm 1\}$ preserves the valuation v. Let $P(\mathbb{Q}) = \left\{ \left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right) \in Sp_{2n}(\mathbb{Q}) | a \in GL_n(\mathbb{Q}) \right\}$. We see easily that $\tau \left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right)^{-1}$ for $\left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right) \in P(\mathbb{Q})^+ := P(\mathbb{Q}) \cap G(\mathbb{Q})^+$ sends q-expansion $\sum_{\xi} a_{\xi} q^{\xi}$ of a modular form to $\sum_{\xi} a_{\xi} \exp(2\pi i \mathrm{Tr}(\xi b d^{-1})) q^{d^{-1}\xi a}$. Thus the image of $P(\mathbb{Q})^+$ in $\overline{\mathcal{E}}(G, X)$ preserves v. For the unipotent radical U of P, the ad-

ditive group $U(\mathbb{Q})$ is dense in $U(\mathbb{A}^{(\infty)})$. Thus $U(\mathbb{A}^{(\infty)}) \subset \mathcal{D}_p$, since $\mathcal{D}_p$ is a closed subgroup. Since $S_1^{(p)}$ is irreducible (if $p \nmid N$), the action of $G_1(\mathbb{A}^{(p\infty)})$ on $\mathbb{Q}^{(p)}(S^{(p)}) = \mathfrak{K}^{G(\mathbb{Z}_p)}$ preserves v restricted to $\mathfrak{K}^{G(\mathbb{Z}_p)}$. Indeed, the action of $G_1(\mathbb{A}^{(p\infty)})$ on $H^0(S^{(p)}_{/\mathcal{W}}, \det(\underline{\omega})^{\otimes k})$ preserves v by the irreducibility of $S^{(p)}_{1/\mathbb{F}}$. By the density theorem Theorem 8.3,

$$D' = \{f \in \bigoplus_{k\geq 0} H^0(S^{(p)}_{/\mathcal{W}}, \det(\underline{\omega})^{\otimes k}) \otimes_{\mathbb{Z}} \mathbb{Q} | v(f) \geq 0\}$$

is dense in the space of p-adic module forms of level p^∞. From this, we conclude that v restricted to $\mathfrak{K}^{U(\mathbb{Z}_p)}$ is preserved by $G_1(\mathbb{A}^{(p\infty)})U(\mathbb{Z}_p)/\{\pm 1\}U(\mathbb{Z}_p)$, which is enough to show the stability of v on $\mathfrak{K}$ under $G_1(\mathbb{A}^{(p\infty)})/\{\pm 1\}$ (because $\tau(U(\mathbb{Z}_p)) \subset \tau(U(\mathbb{A}^{(\infty)}) \subset \mathcal{D}_p$).

The closure of the image of $G_1(\mathbb{A}^{(p\infty)})P(\mathbb{Q})$ in $\overline{\mathcal{E}}(G,X)$ is in $\mathcal{D}_p$. Note that the projection of $G_1(\mathbb{A}^{(p\infty)})P(\mathbb{Q})$ into $G_1(\mathbb{Q}_p)$ contains $P(\mathbb{Q}) \subset G_1(\mathbb{Q}_p)$, which is dense in $P(\mathbb{Q}_p)$. Thus $\mathcal{D}_p$ contains the image under τ of

$$\left\{ \begin{pmatrix} {}^t a^{-1} & 0 \\ 0 & a \end{pmatrix} \in G_1(\mathbb{Z}_p) \right\} \cong GL_n(\mathbb{Z}_p)$$

and $U(\mathbb{Z}_p)$ for the unipotent radical U of P.

The action of $\alpha \in GL_n(\mathbb{Z}_p)$ is given by $(A, \lambda, \eta) \mapsto (A, \lambda, \eta \circ \alpha)$, which induces the action $i_p \mapsto i_p \circ a$ for $\alpha = \left(\begin{smallmatrix} {}^t a^{-1} & 0 \\ 0 & a \end{smallmatrix} \right)$. This shows that $\mathrm{Gal}(C^\circ_\infty / C^\circ_0) = GL_n(\mathbb{Z}_p)$, and hence $T_{K,\alpha} = C_\alpha$ is irreducible over $\overline{\mathbb{F}}_p$. □

Here is the determination of the p-decomposition group $\mathcal{D}_p$ in $\frac{G(\mathbb{A}^{(\infty)})_+}{Z(\mathbb{Q})}$.

Theorem 6.28 *Let $\mathcal{D}_p \subset \mathrm{Aut}(\mathfrak{K})$ be the decomposition group of the valuation v defined in the above proof. The morphism τ in* Theorem 6.27 *induces an isomorphism* $\left\{ g \in G(\mathbb{A}^{(\infty)}) | g_p \in P(\mathbb{Q}_p),\ \nu(g) \in \overline{\mathbb{Q}_p^\times \mathbb{Q}_+^\times} \right\} / Z(\mathbb{Q}) \cong \mathcal{D}_p$.

Proof. We have already shown that $\tau(P(\mathbb{Q})) \subset \mathcal{D}_p$ and $\tau(G_1(\mathbb{A}^{(p\infty)})) \subset \mathcal{D}_p$. Since $\tau \left(\begin{smallmatrix} t\cdot 1_n & 0 \\ 0 & 1_n \end{smallmatrix} \right)$ with $t \in \mathbb{A}^\times$ acts on the q-expansion coefficients at the infinity cusp through the Artin reciprocity law, we find that $\tau \left(\begin{smallmatrix} t\cdot 1_n & 0 \\ 0 & 1_n \end{smallmatrix} \right)$ with $t \in \mathbb{Q}_p^\times$ is in $\mathcal{D}_p$. These elements topologically generate the subgroup of $\overline{\mathcal{E}}(G,X)$ in the corollary. Conversely, if $\tau(g) \in \mathcal{D}_p$ for $g \in G(\mathbb{A}^{(\infty)})$, since $\tau(g)$ acts by $\nu(g)$ on $\mathbb{Q}^{cyc}$, we find that $\nu(g) \in \overline{\mathbb{Q}_p^\times \mathbb{Q}_+^\times}$. Thus modifying g by an element in $P(\mathbb{Q})$, we may assume that $g_p \in G_1(\mathbb{Z}_p)$ for $G_1 = Sp(2n)$. Thus we may assume that $\mathbb{F} = \overline{\mathbb{F}}_p$.

Let $D = \{g_p \mid \tau(g) \in \mathcal{D}_p \text{ and } g_p \in G_1(\mathbb{Z}_p)\}$. By the above argument, we need to show that $D = P^{(1)}(\mathbb{Z}_p) = P(\mathbb{Z}_p) \cap G_1(\mathbb{Z}_p)$. For each $\alpha \in G(\mathbb{Q}_p)$, we can think of $K_\alpha = \alpha G(\mathbb{Z}_p)\alpha^{-1} \cap G(\mathbb{Z}_p)$. Then Sh_{K_α} is a covering of $Sh^{(p)}_{/\mathbb{Q}} = Sh_{G(\mathbb{Z}_p)}$ in two ways. One is coming from $i : Sh_{K_\alpha} \ni [x,g] \mapsto [x,g] \in Sh^{(p)}$ and the other coming from $\alpha : Sh_{K_\alpha} \ni [x,g] \mapsto [x, g\alpha] \in Sh^{(p)}$. In other words, $\mathfrak{M}_{K_\alpha}$ can be considered as a correspondence $[G_1(\mathbb{Z}_p)\alpha_p G_1(\mathbb{Z}_p)]$ inside $\mathfrak{M}^{(p)} \times \mathfrak{M}^{(p)}_{/\mathbb{F}}$ given by the image of $i \times \alpha$. For a p-isogeny $(A, \overline{\lambda}, \eta^{(p)}) \xrightarrow{\xi}$

$(A_\xi, \overline{\lambda}, \eta^{(p)})$ with ordinary A over a base S on which p is nilpotent, choosing ordinary level structure $(\mathbb{Q}_p/\mathbb{Z}_p)^n \cong A[p^\infty]^{et} \cong A_\xi[p^\infty]^{et}$ and its dual inverse $\mu_{p^\infty}^n \cong A[p^\infty]^\circ \cong A_\xi[p^\infty]^\circ$ under the polarization pairing, we have two identifications $\mathrm{Hom}_S(A_\xi[p^\infty]^{et}, \mathbb{Q}_p/\mathbb{Z}_p) \cong \mathrm{Hom}_S(A_\xi[p^\infty]^{et}, \mathbb{Q}_p/\mathbb{Z}_p) \cong \mathbb{Z}_p^n$ and $\mathrm{Hom}_S(A[p^\infty]^\circ, \mu_{p^\infty}) \cong \mathrm{Hom}_S(A_\xi[p^\infty]^{et}, \mu_{p^\infty}) \cong \mathbb{Z}_p^n$, and ξ induces two $n \times n$ matrices $a, d \in M_n(\mathbb{Z}_p)$ (where a sends $\mathrm{Hom}_S(A_\xi[p^\infty]^{et}, \mathbb{Q}_p/\mathbb{Z}_p)$ into $\mathrm{Hom}_S(A_\xi[p^\infty]^{et}, \mathbb{Q}_p/\mathbb{Z}_p)$). The elementary divisor of $\left(\begin{smallmatrix} a & 0 \\ 0 & d \end{smallmatrix}\right)$ is independent of the choice of the ordinary level structures. Then $Sh_{K_\alpha/S}[\frac{1}{H}]/Sh^{(p)}_{K/S}[\frac{1}{H}]$ for the Hasse invariant H classifies all isogenies $(A, \overline{\lambda}, \eta^{(p)}) \xrightarrow{\xi} (A_\xi, \overline{\lambda}, \eta^{(p)})$ (for ordinary abelian schemes A) whose elementary divisor is the same as the elementary divisor of α. We may restrict ourselves to ordinary locus $Sh_{K_\alpha/\overline{\mathbb{F}}_p}[\frac{1}{H}]$ to study irreducible components of $Sh_{K_\alpha/\overline{\mathbb{F}}_p}$ because $Sh_{K_\alpha/\overline{\mathbb{F}}_p}[\frac{1}{H}]$ is Zariski dense in $Sh_{K_\alpha/\overline{\mathbb{F}}_p}$ (e.g., [DAV] V.4.3). We call the pair of elementary divisors of a and d the p-isogeny type of ξ, which gives a finer invariant of p-isogenies than the elementary divisor of α.

This interpretation of Sh_{K_α} (as the moduli of isogenies of elementary divisor type α) also works well over any W-algebra as long as we interpret it as a formal subscheme $\widehat{Sh}_{K_\alpha}$ of the self-product of the formal completion $\widehat{Sh}^{(p)}_K$ of $Sh^{(p)}_K[\frac{1}{E}]$ along its modulo p fiber (for a lift E of the Hasse invariant). The scheme $Sh_{K_\alpha/\mathcal{W}}$ is the algebraization of $\widehat{Sh}_{K_\alpha/W}$ realized as a closed subscheme of $Sh^{(p)} \times_{\mathcal{W}} Sh^{(p)}_{/\mathcal{W}}$ (see [DAV] VII.4) and is finite over $Sh^{(p)}_{/\mathcal{W}}$. We count the number of geometrically irreducible components of $Sh_{K_\alpha/\mathbb{F}} = Sh_{K_\alpha/\mathcal{W}} \otimes_{\mathcal{W}} \mathbb{F}$ over $\mathfrak{M}^{(p)}_{/\mathbb{F}}$. We write $\mathfrak{M}_{\alpha/\mathcal{K}}$ ($\mathcal{K} = Frac(\mathcal{W})$) for the unique geometrically irreducible component of $Sh_{K_\alpha/\mathbb{Q}}$ over $\mathfrak{M}^{(p)}_{/\mathcal{K}}$. Since $\mathrm{Gal}(\mathfrak{M}_{/\overline{\mathbb{Q}}}/\mathfrak{M}^{(p)}_{/\overline{\mathbb{Q}}}) = G_1(\mathbb{Z}_p)$, by Hilbert's theory of decomposition of primes ([BCM] V.5.2–3), the irreducible components modulo $\mathfrak{m}_{\mathcal{W}}$ are in bijection with double cosets $K_\alpha^{(1)} \backslash G_1(\mathbb{Z}_p)/D$ for $K_\alpha^{(1)} = K_\alpha \cap G_1(\mathbb{Z}_p)$, since irreducible components correspond to minimal primes of the integral closure in the function field of $\mathfrak{M}_{K_\alpha/\mathcal{K}}$ of the valuation ring of v in the function field of $\mathfrak{M}^{(p)}_{/\mathcal{K}}$. Let $\alpha_m = \left(\begin{smallmatrix} 1_n & 0 \\ 0 & p^m 1_n \end{smallmatrix}\right)$. Thus if we can prove

$$|K_m^{(1)} \backslash G_1(\mathbb{Z}_p)/D| \geq |K_m^{(1)} \backslash G_1(\mathbb{Z}_p)/K_m^{(1)}|$$

for all m, we conclude $D = \bigcap_m K_m^{(1)} = P^{(1)}(\mathbb{Z}_p)$, since $P^{(1)}(\mathbb{Z}_p) \subset D$. We let $G(\mathbb{Z}_p)$ act on $L_m = (\mathbb{Z}/p^m\mathbb{Z})^{2n}$ by matrix multiplication. We consider the set $\mathcal{L}_m$ of maximally isotropic free $\mathbb{Z}/p^m\mathbb{Z}$-submodules of L_m; so,

$$\mathcal{L}_m = \left\{ V \subset L_m \;\middle|\; V \cong (\mathbb{Z}/p^m\mathbb{Z})^n \text{ and } {}^t v \cdot J_n v' = 0 \text{ for all } v, v' \in V \right\}.$$

Then $G_1(\mathbb{Z}_p)$ acts on $\mathcal{L}_m$ transitively. We write V_m for the element of $\mathcal{L}_m$ generated by the first n standard basis of $(\mathbb{Z}/p^m\mathbb{Z})^{2n}$. Then $K_m^{(1)} \subset K^{(1)}$ is the

stabilizer of V_m in $\mathcal{L}_m$. Identifying $\mathcal{L}_m$ with the subgroups of $A[p^m]$ of a generic p-ordinary member $A_{/W}$ of the universal abelian scheme $\mathcal{A}_{/\mathfrak{M}^{(p)}}$, $\mathfrak{M}_m = \mathfrak{M}_{K_m}$ classifies isogenies $\phi : A \to A/V$ whose kernel is free of rank n over $\mathbb{Z}/p^m\mathbb{Z}$ generically. On the special fiber of characteristic p of A, this kernel $V = \mathrm{Ker}(\phi)$ degenerates into a $\mathbb{Z}/p^m\mathbb{Z}$-submodule of $A[p^m](\mathbb{F}) \cong (\mathbb{Z}/p^m\mathbb{Z})^n$. This image is determined up to isomorphisms by the intersection $V \cap A[p^m]^\circ(\overline{\mathbb{Q}}_p)$ for the connected component $A[p^m]^\circ$ over $\mathcal{W}$, which also determines the p-isogeny type of ϕ. In [Sh1] Theorem 5, it is proven (under some assumptions later removed in [DAV] VII.4) that each irreducible component of $\mathfrak{M}_{m/\mathbb{F}}$ is determined by the characteristic 0 intersection $V \cap A[p^m]^\circ(\overline{\mathbb{Q}}_p)$. We only need a weaker assertion that the p-isogeny type only depends on components. Indeed, the p-isogeny type is locally constant on $Sh_{K_m/\overline{\mathbb{F}}_p}[\frac{1}{H}]$ by the Serre–Tate deformation theory (see Theorem 8.9 and the argument around (8.25)). Thus $|K_m^{(1)} \backslash G_1(\mathbb{Z}_p)/D| \geq |\mathcal{L}_m/\sim|$, where $V \sim V'$ if $V \cap (A[p^m]^\circ(\overline{\mathbb{Q}}_p)) = V' \cap (A[p^m]^\circ(\overline{\mathbb{Q}}_p))$. It is easy to show that the number $|\mathcal{L}_m/\sim|$ is equal to $|K_m^{(1)} \backslash G_1(\mathbb{Z}_p)/K_m^{(1)}|$, which shows the desired assertion (and also the bijection between $\mathcal{L}_m/\sim$ and irreducible components of $\mathfrak{M}_{K_m/\mathbb{F}}$). We refer the reader to the book of Faltings and Chai [DAV] VII.4 for more detailed information on the mod p correspondence $[G_1(\mathbb{Z}_p)\alpha_p G_1(\mathbb{Z}_p)]$. □

7

Shimura Varieties

In this chapter, we describe a construction of Shimura varieties of PEL type following [Sh5], [D2], and [Ko]. The construction is given in the first section. The last theorem (Theorem 7.7) follows basically from the discussion of the first section if we limit our groups G to those described in the first section; however, the theorem is formulated for more general groups.

Shimura originally constructed canonical models in the 1950s to the 1960s as a tower of quasi-projective geometrically connected varieties (over a tower of canonical abelian extensions of the reflex field) with a specific reciprocity law at special algebraic points (in the case of Shimura varieties of PEL type, they are called *CM points* carrying abelian varieties of CM type; [Sh5]). His theory includes interesting cases of canonical models of non-PEL type (e.g., Shimura curves $Y_0^D(\mathfrak{N})$ for totally real fields $F \neq \mathbb{Q}$ studied in Section 4.3.3), but in this book, we restrict ourselves to the case where we have a canonical family of abelian varieties over the canonical model (so, the construction of the models is easier, as was basically done in [Sh3]).

Deligne reformulated Shimura's tower as a projective limit of (possibly nonconnected) models over the reflex field (incorporating the theory of motives in its scope). We follow Deligne's treatment in order to avoid the definition of the canonical fields of definition of the connected components, although the global reciprocity law (Theorem 7.7) essentially contains this information.

Kottwitz extended Deligne's definition of Shimura varieties of PEL type to a projective limit of schemes over a valuation ring of mixed characteristic $(0, p)$, when the level is prime to p. Since we are interested in the formal completion at p of the Kottwitz model (and an analogue of the Igusa tower over the Kottwitz model), what we use most in the following chapter for p-adic automorphic forms is Kottwitz's formulation.

Under the notation we introduced for $GL(2)$ and $GSp(2n)$, roughly speaking, the canonical model of Deligne is given by $Sh(G, X)$, the model of Kottwitz is given by $Sh^{(p)}(G, X)$, and the model of Shimura is the geometrically irreducible component of $Sh(G, X)$ which is the image of $X^+ \times 1$.

7.1 PEL Moduli Varieties

We construct the moduli of abelian schemes with specific endomorphism algebra, which is a simple algebra. We fix a rational prime p.

7.1.1 Polarization, Endomorphism, and Lattice

Let D be a finite-dimensional simple $\mathbb{Q}$-algebra with center F. We write $O = O_F$ for the integer ring of F. Let $\mathcal{S}$ be a set of prime ideals of O over p. We always assume

(unr) We have an isomorphism $D_p = D \otimes_{\mathbb{Q}} \mathbb{Q}_p \cong \bigoplus_{\mathfrak{p}\in\mathcal{S}} M_d(F_{\mathfrak{p}})$ and $F_{\mathfrak{p}}/\mathbb{Q}_p$ is unramified for all $\mathfrak{p} \in \mathcal{S}$.

Let $*$ be an involution on D that satisfies $\mathrm{Tr}_{D/\mathbb{Q}}(xx^*) > 0$ for all $0 \neq x \in D_\infty = D \otimes_{\mathbb{Q}} \mathbb{R}$ (a *positive involution*). Here $\mathrm{Tr}_{D/\mathbb{Q}} : D \to \mathbb{Q}$ is the reduced trace Tr of D over F composed with the field trace $\mathrm{Tr}_{F/\mathbb{Q}}$. We call such an involution a *positive* involution. We define a subfield F_0 of F by the field fixed by the involution $*$. Then F_0 has to be totally real because $\mathrm{Tr}_{F_0/\mathbb{Q}}(x^2) = \mathrm{Tr}_{F_0/\mathbb{Q}}(xx^*) > 0$ for all $x \in F_0^\times$ is true only for totally real fields. Then we conclude either $F = F_0$ or F is a totally imaginary quadratic extension of F_0 (a CM field). If $F \neq F_0$, the involution $*$ coincides on F with "complex conjugation" which is the generator of $\mathrm{Gal}(F/F_0)$. The involution $*$ is called an involution of *the first kind* if $F = F_0$, and otherwise it is called an involution of *the second kind*. We fix a maximal order O_D of D stable under $*$ (see [BNT] V.1 and X.1 for orders). It is an easy exercise to show that we can choose a maximal order O_D stable under any given involution of D. We assume that the isomorphism in (unr) induces

$$O_{D,p} = O_D \otimes_{\mathbb{Z}} \mathbb{Z}_p \cong \bigoplus_{\mathfrak{p}\in\mathcal{S}} M_d(O_{F,\mathfrak{p}}).$$

We fix a left D-module V of finite type and assume that we have a nondegenerate alternating form $\langle\ ,\ \rangle : V \times V \to \mathbb{Q}$ such that $\langle bv, w\rangle = \langle v, b^*w\rangle$ for all $b \in D$. Write $V_p = V \otimes_{\mathbb{Q}} \mathbb{Q}_p$ and $V_\infty = V \otimes_{\mathbb{Q}} \mathbb{R}$. We also assume having an O_D-submodule $L \subset V$ of finite type such that

(L1) $L \otimes_{\mathbb{Z}} \mathbb{Q} = V$;
(L2) $\langle\ ,\ \rangle$ induces $\mathrm{Hom}_{\mathbb{Z}_p}(L_p, \mathbb{Z}_p) \cong L_p$, where $L_p = L \otimes_{\mathbb{Z}} \mathbb{Z}_p$.

Put $C = \mathrm{End}_D(V)$, which is a semi-simple $\mathbb{Q}$-algebra with involution denoted by ι (which is often not positive) given by $\langle cv, w\rangle = \langle v, c^\iota w\rangle$. Then we define algebraic $\mathbb{Q}$-groups G, GU, and U by

$$G(R) = \left\{x \in C \otimes_{\mathbb{Q}} R \middle| xx^\iota \in R^\times\right\},\quad U(R) = \left\{x \in G(R) \middle| xx^\iota = 1\right\},$$
$$\text{and}\quad GU(R) = \left\{x \in C \otimes_{\mathbb{Q}} R \middle| xx^\iota \in (F_0 \otimes_{\mathbb{Q}} R)^\times\right\}. \quad (7.1)$$

As before, we write G_1 for the derived group of G (and GU), and assume

(sc) *G_1 is simply connected with noncompact $G_1(\mathbb{R})$.*

By this assumption, we have the strong approximation theorem valid for G_1, that is, $G_1(\mathbb{Q})$ is dense in $G_1(\mathbb{A}^{(\infty)})$ (see [Kn]). This type of group covers all the cases of Shimura varieties we have studied so far. If we take D to be a totally real field F with the identity involution $*$, $V = F^2$, $L = O \oplus \mathfrak{c}^*$, and $\langle (x,x'),(y,y')\rangle = xy' - yx'$, we have $C = M_2(F)$ with involution $x \mapsto x^\iota = \det(x)x^{-1}$ and therefore $GU = \mathrm{Res}_{F/\mathbb{Q}}GL(2)$ and $G_1 = U = \mathrm{Res}_{F/\mathbb{Q}}SL(2)$; so, we are in the Hilbert modular case. If we take $D = \mathbb{Q}$ with the identity involution, $V = \mathbb{Q}^{2g}$ and $L = \mathbb{Z}^{2g}$ with $\langle x,y\rangle = {}^t x \cdot J_g y$, we have $C = M_{2g}(\mathbb{Q})$ with involution $x \mapsto J_g {}^t x \cdot J_g^{-1}$ and therefore $G = GSp(2g)$ and $G_1 = U = Sp(2g)$; so, we are in the Siegel modular case. If $*$ is of the second kind, U is not equal to the derived group $G_1 = GU^{der} = G^{der}$, and G_1 is often written as SU (a *special unitary* group). We have the inclusion $G_1 = SU \subset U \subset G \subset GU$ and exact sequences of $fppf$-sheaves (see Section 8.2.1 for $fppf$-sheaves),

$$1 \to U \to G \xrightarrow{\nu} \mathbb{G}_m \to 1,\ 1 \to U \to GU \xrightarrow{\nu} \mathrm{Res}_{F_0/\mathbb{Q}}\mathbb{G}_m \to 1,$$
$$1 \to SU \to U \xrightarrow{\det} \mathrm{Res}_{F_0/\mathbb{Q}}\mathrm{Ker}(Norm : \mathrm{Res}_{F/F_0}\mathbb{G}_m \to \mathbb{G}_{m/F_0}) \to 1,$$
$$\text{and } 1 \to SU \to GU \xrightarrow{\det \times \nu} \mathrm{Res}_{F/\mathbb{Q}}\mathbb{G}_m \times \mathrm{Res}_{F_0/\mathbb{Q}}\mathbb{G}_m, \quad (7.2)$$

where $\nu(x) = xx^\iota$.

In order to bring a complex structure into the real vector space $V_\infty = V \otimes_{\mathbb{Q}} \mathbb{R}$, we use an $\mathbb{R}$-algebra homomorphism $h : \mathbb{C} \hookrightarrow C_\infty = C \otimes_{\mathbb{Q}} \mathbb{R}$ with $h(\overline{z}) = h(z)^\iota$. We call such an algebra homomorphism *ι-homomorphism*. Then $h(i)^\iota = -h(i)$ for $i = \sqrt{-1}$ and hence $x^\rho = h(i)^{-1}x^\iota h(i)$ is an involution of C_∞. We suppose

(pos) The symmetric real bilinear form $(v,w) \mapsto \langle v, h(i)w\rangle$ on V_∞ is positive-definite.

Thus we have $0 < (xv,xv) = (v,(xx^\rho)v)$ for all $0 \neq v \in V_\infty$ and $x \in C_\infty$, and hence xx^ρ only has positive eigenvalues. In particular, ρ is a positive involution of C (i.e., $\mathrm{Tr}_{C/\mathbb{Q}}(xx^\rho) > 0$ unless $x = 0$). Such an $h : \mathbb{C} \to C_\infty$ induces a homomorphism $h : \mathbb{S} := \mathrm{Res}_{\mathbb{C}/\mathbb{R}}\mathbb{G}_m \to G_{/\mathbb{R}}$ of algebraic groups over $\mathbb{R}$. We can easily verify that this homomorphism satisfies the conditions (D1–2) in the following section, and to have an algebra homomorphism $h : \mathbb{C} \to C_\infty$ as above is equivalent to having a group homomorphism $h : \mathbb{S}_{/\mathbb{R}} \to G_{/\mathbb{R}}$ satisfying (D1–2). Any two homomorphisms satisfying (pos) are conjugates under $G_1(\mathbb{R})^+$ (see Section 7.1.5). Fix one such $h := h_{\mathbf{0}} : \mathbb{S} \to G$ satisfying (D1–2), and define $X = X_G$ for the collection of all conjugates of $h_{\mathbf{0}}$ under $G(\mathbb{R})$. On X, $G(\mathbb{R})$ acts by conjugation (from the left), and the stabilizer $C_{\mathbf{0}} \subset G(\mathbb{R})$ of $h_{\mathbf{0}}$ is a maximal compact subgroup of $G(\mathbb{R})$ modulo center by (pos). We can also think of the set X_{GU} of all conjugates of $h_{\mathbf{0}}$ under $GU(\mathbb{R})$.

The connected components X_G^+ and X_{GU}^+ containing h_0 of these two spaces are identical.

Since $h : \mathbb{C} \to C_\infty$ is an $\mathbb{R}$-algebra homomorphism, we can split $V_\mathbb{C} = V \otimes_\mathbb{Q} \mathbb{C}$ into the direct sum of eigenspaces $V_\mathbb{C} = V_1 \oplus V_2$ so that $h(z)$ acts on V_1 (resp. V_2) through multiplication by z (resp. $\overline{z}$); thereby, we get a complex vector space structure on V_∞ by the projection $V_\infty \cong V_1$. Since $h(\mathbb{C}) \subset C_\infty$, $h(z)$ commutes with the action of D; so, V_j is stable under the action of $D_\mathbb{C} = D \otimes_\mathbb{Q} \mathbb{C}$. We get the representation $\rho_1 : D \hookrightarrow \mathrm{End}_\mathbb{C}(V_1)$. We define E for the subfield of $\mathbb{C}$ fixed by $\left\{\sigma \in \mathrm{Aut}(\mathbb{C}) \middle| \rho_1^\sigma \cong \rho_1\right\}$. If $h'(z) = g \cdot h(z) g^{-1}$ for $g \in GU(\mathbb{R})$, h' induces a similar decomposition $V_\mathbb{C} = V_1' \oplus V_2'$, and g induces a D-linear isomorphism between V_1 and V_1'; so, E is independent of the choice of h' in the $GU(\mathbb{R})$-conjugacy class of h. This field E is called the *reflex field* of (G, X) (and is a canonical field of definition of our canonical models of the Shimura variety). By the positivity (pos), the quotient complex torus $V_\infty/L = V_1/L$ has a Riemann form induced by $\langle \cdot, \cdot \rangle$. The theta functions with respect to the associated Hermitian form $(\cdot, \cdot)$ in (pos) gives global sections of an ample line bundle (e.g., [ABV] Chapter I) on V_1/L and hence embed V_1/L into a projective space over $\mathbb{C}$. The embedded image is the analytic space associated with an abelian variety $A_{h/\mathbb{C}}$. Multiplication by $b \in O_D$ on V_1/L induces an embedding $i : D \hookrightarrow \mathrm{End}^\mathbb{Q}(A_{h/\mathbb{C}})$. Since $\langle bx, y \rangle = \langle x, b^* y \rangle$, $i(b) \mapsto i(b^*)$ is induced by the Rosati involution of the polarization of the Riemann form $\langle \cdot, \cdot \rangle$. The representation ρ_1 is given by the action of D on the tangent space $Lie(A_h) = V_1$ at the origin of A_h. Since A_h is projective, the field of definition of the abelian variety A_h is a field of finite type over $\mathbb{Q}$. By Corollary 6.11, the couple $(A_h, \iota : D \hookrightarrow \mathrm{End}^\mathbb{Q}(A_h))$ is defined over a field of finite type over $\mathbb{Q}$. By definition, the field E is the field of rationality of the representation of D on $Lie(A_h)$; so, the field of definition of (A_h, ι) always contains this field E. It would then be natural to expect that the moduli variety of triples (A, λ, ι) for an abelian variety A with D-linear isomorphism $Lie(A) \cong V_1$ is defined over E.

Since the isomorphism class of ρ_1 is determined by $\mathrm{Tr}(\rho_1)$ (see [MFG] Proposition 2.9), E is generated over $\mathbb{Q}$ by $\mathrm{Tr}(\rho_1(b))$ for all $b \in D$. The field E is a subfield of the Galois closure F^{gal} of $F/\mathbb{Q}$, and E is a finite extension of $\mathbb{Q}$. Moreover, under (unr), p is unramified in $E/\mathbb{Q}$. We write O_E for the integer ring of E. Let $\mathbb{Z}_{(p)} = \mathbb{Z}_p \cap \mathbb{Q}$ and put $O_{D,(p)} = O \otimes_\mathbb{Z} \mathbb{Z}_{(p)}$, $O_{(p)} = O \otimes_\mathbb{Z} \mathbb{Z}_{(p)}$, and $\mathcal{V} = O_E \otimes_\mathbb{Z} \mathbb{Z}_{(p)}$. The ring $\mathcal{V}$ is a semi-local Dedekind domain with quotient field E in which p is unramified by (unr).

Let $K^{(p)}$ be an open compact subgroup of $G(\mathbb{A}^{(p\infty)})$. We study the classification problem of the following quadruples $(A, \lambda, i, \overline{\eta}^{(p)})_{/S}$: A is a (projective) abelian scheme over a base S, ${}^tA = \mathrm{Pic}^0_{A/S}(A)$ is the dual abelian scheme of A, $\lambda : A \to {}^tA$ is an isogeny with degree prime to p (prime-to-p isogeny) fiber-by-fiber geometrically induced from an ample divisor (polarization), $i : O_{D,(p)} \hookrightarrow \mathrm{End}_S^{\mathbb{Z}_{(p)}}(A) = \mathrm{End}_S(A) \otimes_\mathbb{Z} \mathbb{Z}_{(p)}$ is a $\mathbb{Z}_{(p)}$-algebra embedding (taking 1 to the identity of A) with $\lambda \circ i(\alpha^*) = {}^ti(\alpha) \circ \lambda$ for all $\alpha \in O_D$,

and $\overline{\eta}^{(p)}$ is the level $K^{(p)}$-structure. In other words, regarding ${}^tA$ as a left O_D-module by $O_D \ni b \mapsto {}^ti({}^*) \in \mathrm{End}({}^tA)$, λ is D-linear. Hereafter we call λ *D-linear* in this sense. The base scheme S is assumed to be a scheme over $Spec(\mathcal{V})$.

We clarify the meaning of the level $K^{(p)}$-structure $\overline{\eta}^{(p)}$. Fix a base (geometric) point $s \in S$ and write A_s for the fiber of A at s. We consider the Tate module $\mathcal{T}(A_s) = \varprojlim_N A[N](k(s))$ and $V^{(p)}(A) = \mathcal{T}(A) \otimes_{\mathbb{Z}} \mathbb{A}^{(p\infty)}$, where N runs over all positive integers ordered by divisibility. The prime-to-p level structure $\eta^{(p)} : V(\mathbb{A}^{(p\infty)}) = V \otimes_{\mathbb{Q}} \mathbb{A}^{(p\infty)} \cong V^{(p)}(A_s)$ is an O_D-linear isomorphism. The duality pairing $e_N : A[N] \times {}^tA[N] \to \mu_N$ composed with λ gives, after taking the limit with respect to N, an alternating form $\langle \cdot, \cdot \rangle_\lambda : V^{(p)}(A_s) \times V^{(p)}(A_s) \to \mathbb{A}^{(p\infty)}(1) := \mathbb{Q} \otimes_{\mathbb{Z}} \varprojlim_{p \nmid N} \mu_N$ (see (P1–3) in Section 8.2.3). Hereafter we require that $\eta^{(p)}$ send the alternating form $\langle \cdot, \cdot \rangle$ to $\langle \cdot, \cdot \rangle_\lambda$ up to the scalar multiple by $(\mathbb{A}^{(p\infty)})^\times$. This is possible, because $\mathbb{A}^{(p\infty)}(1) \cong \mathbb{A}^{(p\infty)}$ up to the scalar multiple by $(\mathbb{A}^{(p\infty)})^\times$. Then $\eta^{(p)}$ is required to be an isomorphism of skew Hermitian D-modules with respect to the pairing $\langle \cdot, \cdot \rangle_\lambda$ on $\mathcal{T}(A)$.

The algebraic fundamental group $\pi_1(S, s)$ (cf. Section 4.4) acts on $V^{(p)}(A_s)$ leaving the skew Hermitian form $\langle \cdot, \cdot \rangle_\lambda$ stable up to the scalar (because it keeps the Weil e_N-pairing; see [ABV] Section 20). We write $\overline{\eta}^{(p)}$ for the orbit $\eta^{(p)} \circ K^{(p)}$ and suppose that $\sigma \circ \overline{\eta}^{(p)} = \overline{\eta}^{(p)}$ for all $\sigma \in \pi_1(S, s)$ (this is a way of describing that the level structure $\eta^{(p)}$ is defined over S). Even if we change the point $s \in S$, everything will be conjugated by an isomorphism (as seen in Section 6.4.1); so, the definition does not depend on the choice of s as long as S is connected. For nonconnected S, we choose one geometric point at each connected component.

As examples of $K^{(p)}$ and open compact subgroups of $G(\mathbb{A}^{(\infty)})$, we could offer the following subgroups for $\widehat{L} = L \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}$,

$$\widehat{\Gamma} = \left\{ x \in G(\mathbb{A}^{(\infty)}) \middle| x\widehat{L} = \widehat{L} \right\}, \ \widehat{\Gamma}^{(p)} = \left\{ x \in \widehat{\Gamma} \middle| x_p = 1 \right\};$$

$$\widehat{\Gamma}(N) = \left\{ x \in \widehat{\Gamma} \middle| x\ell \equiv \ell \mod N\widehat{L} \ \forall \ell \in \widehat{L} \text{ for an integer } p \nmid N > 0 \right\}.$$

An open compact subgroup $K = K_p \times K^{(p)}$ of $G(\mathbb{A}^{(\infty)})$ is called *maximal* at p if $K_p = \widehat{\Gamma}_p$ for the p-component $\widehat{\Gamma}_p$ of $\widehat{\Gamma}$ as above.

Since a maximal compact subgroup of $GL_D(V(\mathbb{A}^{(\infty)})) = \mathrm{Aut}_{D_{\mathbb{A}}}(V(\mathbb{A}^{(\infty)}))$ is the stabilizer of a $\widehat{\mathbb{Z}}$-lattice in $V(\mathbb{A}^{(\infty)})$ stable under a maximal order O_D, we find a lattice L with $L \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}^{(p)}$ stable under $K^{(p)}$, where $\widehat{\mathbb{Z}}^{(p)} = \prod_{\ell \neq p} \mathbb{Z}_\ell$ for ℓ running through all primes different from p. Changing L by a sublattice of p-power index if necessary, we may assume that L satisfies the conditions (L1–2). We call a quadruple $\underline{A}_{/S} = (A, \lambda, i, \overline{\eta}^{(p)})_{/S}$ isomorphic to $\underline{A}'_{/S} = (A', \lambda', i', \overline{\eta}'^{(p)})_{/S}$ if we have an O_D-linear isogeny $\phi : A \to A'$ defined over S such that $p \nmid \deg(\phi)$, $\phi^*\lambda' = {}^t\phi \circ \lambda' \circ \phi = c\lambda$ with $c \in \mathbb{Z}^\times_{(p)}$, $\phi \circ i \circ \phi^{-1} = i'$, and $\overline{\eta}'^{(p)} = \phi \circ \overline{\eta}^{(p)}$. Thus ϕ brings the prime-to-p polarization class $\overline{\lambda}' = \{c\lambda' | c \in$

$\mathbb{Z}_{(p)}^{\times}\}$ of λ' to the class $\overline{\lambda}$ of λ: $\phi^*\overline{\lambda}' = \overline{\lambda}$. In this case, we write $A \approx A'$. We write $A \cong A'$ if the isogeny is an isomorphism of abelian schemes; that is, $\deg(\phi) = 1$.

We take the fibered category $\mathcal{C} = \mathcal{C}_{D,V}$ of the quadruples $(A, \lambda, i, \overline{\eta}^{(p)})_{/S}$ over the category $\mathcal{V}$-SCH of $\mathcal{V}$-schemes and define

$$\begin{aligned}&\mathrm{Hom}_{\mathcal{C}_{/S}}((A,\lambda,i,\overline{\eta}^{(p)})_{/S},(A',\lambda',i',\overline{\eta}'^{(p)})_{/S})\\ &=\left\{\phi\in \mathrm{Hom}_S(A,A')\otimes_{\mathbb{Z}}\mathbb{Z}_{(p)}\,\middle|\, \begin{matrix}{}^t\phi\circ\lambda'\circ\phi = c\lambda \ \text{ with } \ 0<c\in\mathbb{Z}_{(p)}^{\times},\\ \phi\circ i = i'\circ\phi \ \text{ and } \ \overline{\eta}'^{(p)} = \phi\circ\overline{\eta}^{(p)}\end{matrix}\right\}. \qquad (7.3)\end{aligned}$$

We consider the functor $\mathcal{E}_K^{(p)} : S_0$-$SCH \to SETS$ given by

$$\mathcal{E}^{(p)}(S) = \mathcal{E}_K^{(p)}(S) = \left\{\underline{A}_{/S} = (A,\overline{\lambda},i,\overline{\eta}^{(p)})_{/S}\,\middle|\, A \text{ satisfies (det) below}\right\}/\approx .$$

Here the determinant condition (det) is given as follows. Recall $\mathcal{V} = O_E \otimes_{\mathbb{Z}} \mathbb{Z}_{(p)}$ for the integer ring O_E of E. As already remarked, $\mathcal{V}$ is a semi-local Dedekind domain with quotient field E. The condition (unr) guarantees that $O_D \otimes \mathcal{V}/\mathfrak{m}$ is semi-simple for all maximal ideals $\mathfrak{m} \subset \mathcal{V}$. We choose a $\mathbb{Z}_{(p)}$-base $\{\alpha_\jmath\}_{1\le\jmath\le r}$ of $O_{D,(p)}$ and consider a homogeneous polynomial

$$f(t_1,\dots,t_r) = \det(\alpha_1 t_1 + \cdots + \alpha_t t_r|_{V_1}).$$

Then $f(t) \in \mathcal{V}[t_1,\dots,t_r]$, and the coefficients of $f(t)$ generate E over $\mathbb{Q}$. Here O_E is the integer ring of E. Since $A_{/S}$ is a group scheme, its tangent space at the zero section has a Lie algebra structure over $\mathcal{O}_S$. We write $Lie(A)$ for this Lie algebra. Since A is smooth over S, $Lie(A)$ is a locally free $\mathcal{O}_S$-module of rank $\dim_S A$ (see [NMD] Proposition 2.2.5). In our case, for a given quadruple $\underline{A} = (A,\lambda,i,\overline{\eta}^{(p)})_{/S}$, the Lie algebra $Lie(A)$ of A over $\mathcal{O}_S$ is an $O_{D,(p)} \otimes_{\mathbb{Z}} \mathcal{O}_S$-module via i. Since $Lie(A)$ is locally free of rank $\dim_S A$ over $\mathcal{O}_S$, we can think of $g(t_1,\dots,t_r) = \det(\alpha_1 t_1 + \cdots + \alpha_r t_r|_{Lie(A)}) \in \mathcal{O}_S[t_1,\dots,t_r]$. We impose

$$j(f(t_1,\dots,t_r)) = g(t_1,\dots,t_r) \ \text{ in } \mathcal{O}_S[t], \qquad \text{(det)}$$

where $j : \mathcal{V} \to \mathcal{O}_S$ is the structure homomorphism. Since $f(t)$ contains as coefficients all data of the characteristic polynomial of $\rho_1(\alpha)$, this condition is equivalent to (det) in Section 6.1.6. Over a field of characteristic 0, one can characterize representations of a semi-simple algebra by their trace, but over a general base, we need (as explained in Section 6.1.6 around the condition (det) there) the entire characteristic polynomials to determine a given representation; so, the determinant has to be fixed as above.

In the rest of this section, we give a sketch of a proof of the following theorem.

Theorem 7.1 *Let the notation and the assumptions be as above. If the open subgroup K of $G(\mathbb{A}^{(\infty)})$ maximal at p is sufficiently small, the functor $\mathcal{E}_K^{(p)}$*

is representable by a quasi-projective smooth scheme $Sh_K^{(p)}$ *over* $\mathcal{V}$. *For any* K *maximal at* p, *the coarse moduli scheme* $Sh_K^{(p)}$ *of* $\mathcal{E}_K^{(p)}$ *exists as a quasi-projective scheme of finite type over* $\mathcal{V}$. *If* D *is a division algebra and* $V = D$, $Sh_{K/\mathcal{V}}^{(p)}$ *is projective over* $\mathcal{V}$.

Even if $D = F$ and $V = F^n$ with $n > 1$, $Sh_K^{(p)}$ is projective if the Hermitian pairing $\langle \cdot, \cdot \rangle$ is anisotropic (i.e., $\langle v, v \rangle = 0 \Leftrightarrow v = 0$ for $v \in V$), because the polarization pairing has to have an isotropic subspace given by the toric part of the semi-abelian scheme if the universal abelian variety degenerates into a semi-abelian scheme.

The generic fiber of the variety $Sh_{K/E}^{(p)}$ actually could be a finite disjoint union of Shimura varieties $Sh^{(p)}(G, X_G)$ associated with different pairs (G, X_G) in the sense of Deligne, depending on the choice of $(D, V, \langle \cdot, \cdot \rangle)$ and G, because our data only determine the Hermitian form $\langle \cdot, \cdot \rangle$ over $D_{\mathbb{A}}^{(p\infty)}$ not over D; so, if the Hasse principle fails for such Hermitian forms (i.e., the isomorphism class of $(V_{\mathbb{A}^{(p\infty)}}, \langle \cdot, \cdot \rangle)$ over $D_{\mathbb{A}}^{(p\infty)}$ does not determine uniquely the isomorphism class $(V, \langle \cdot, \cdot \rangle)$ over D), there could be different groups $G_{/\mathbb{Q}}$ associated with the functor $\mathcal{E}_K^{(p)}$. This delicate point is studied later in Section 7.1.5.

7.1.2 Construction of the Moduli

If two abelian schemes A and A' over S are isogenous by a prime-to-p isogeny over S, we write $A \approx A'$, abusing our notation slightly. If $K^{(p)}$ is sufficiently small so that $\mathrm{Aut}_S(A, \lambda, i, \overline{\eta}^{(p)}) = \{1_A\}$, the prime-to-$p$ isogeny giving the isomorphism $A \approx A'$ in the definition of $\mathcal{E}^{(p)}$ can be taken (as explained in Section 4.2.1) to be an isomorphism by changing A' in the isogeny class under "$\approx$" (and insisting $\eta^{(p)}(\widehat{L}^{(p)}) = \mathcal{T}(A) \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}^{(p)}$; cf. [D1] 4.10). Therefore we have $\mathcal{E}_K^{(p)}(S) \cong \mathcal{E}_K(S)$ for another functor $\mathcal{E}_K : SCH_{/\mathcal{V}} \to SETS$ given by

$$\mathcal{E}_K(S) = \left\{ \underline{A}_{/S} \middle| \ A \approx \exists A' \in \mathcal{E}_K^{(p)}(S) \text{ and } \eta^{(p)}(\widehat{L}^{(p)}) = \mathcal{T}(A) \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}^{(p)} \right\} / \cong .$$

Here $\underline{A}_{/S}$ is an abbreviated form of $(A, \lambda, i, \overline{\eta}^{(p)})_{/S}$. Under this setting, we change the morphism set of $\mathcal{C}_{/S}$ from $\mathrm{Hom}_{\mathcal{C}_{/S}}$ to $\mathrm{Isom}_{\mathcal{C}_{/S}}$:

$$\mathrm{Isom}_{\mathcal{C}_{/S}}((A, \lambda, i, \overline{\eta}^{(p)})_{/S}, (A', \lambda', i', \overline{\eta}'^{(p)})_{/S}) = \left\{ \phi \in \mathrm{Isom}_S(A, A') \middle| \begin{matrix} {}^t\phi \circ \lambda' \circ \phi = c\lambda \ \text{ with } \ c \in \mathbb{Z}_{(p)}^{\times}, \\ \phi \circ i = i' \circ \phi \ \text{ and } \ \overline{\eta}'^{(p)} = \phi \circ \overline{\eta}^{(p)} \end{matrix} \right\}. \quad (7.4)$$

Then we claim that the fibered category $\mathcal{C}$ over $\mathcal{V}$-SCH is an algebraic stack for any given $K^{(p)}$ and is an algebraic space if $\mathrm{Aut}_{\mathcal{C}_{/S}}(A) = \{1_A\}$ for all objects A of $\mathcal{C}_{/S}$. See [DM] and [DAV] I.4.6–9 for the definitions of algebraic

stacks, but this statement is basically equivalent to the fact that $\mathcal{E}_K^{(p)}$ has a coarse moduli scheme over $\mathcal{V}$ (compatible with base change). Indeed, if $K^{(p)}$ is sufficiently small, $\mathcal{E}_K^{(p)}$ is representable by a quasi-projective scheme over $\mathcal{V}$ as we show.

Forgetting D-linearity of $\eta^{(p)}$ and restricting i to $\mathbb{Q}$, we have a functor from $\mathcal{C}_{D,V}$ into the fibered category $\mathcal{C}_{\mathbb{Q},V}$ of $(A, \lambda, i, \overline{\eta}^{(p)})$ for $D = \mathbb{Q}$ for a suitable choice of an open compact subgroup $GSp(2d)(\mathbb{A}^{(p\infty)})$ ($d = \dim A = \frac{1}{2}\dim_{\mathbb{Q}} V$). As we have shown in Theorem 6.20, this fibered category is proven to be an algebraic stack and is representable by a quasi-projective scheme $M = M_{\mathbb{Q}/\mathcal{V}}$ by Mumford (see also [GIT], [Sh3], [CSM], and [DAV]).

We now supplement the above outline with details. We show that we can replace $\approx$ by $\cong$ in the definition of the functor $\mathcal{E}_K^{(p)}$, imposing an additional condition. Let $\underline{A}_{/S} = (A, \lambda, i, \overline{\eta}^{(p)})_{/S}$. Then $\eta^{(p)}$ induces $V(\mathbb{A}^{(p\infty)}) \cong V^{(p\infty)}(A)$ and, therefore, we have $\eta^{(p)}(\widehat{L}^{(p)}) \subset V^{(p\infty)}(A)$ for $\widehat{L}^{(p)} = L \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}^{(p)}$.

If $\eta^{(p)}(\widehat{L}^{(p)}) \subset \mathcal{T}^{(p)}(A)$, the cokernel is an étale group subscheme $C \subset A_{/S}$ locally free over S of rank prime to p. Make the quotient abelian scheme $A' = A/C$ over S (see [ABV] Section 12 and [GME] Proposition 1.8.4), and then we have a prime-to-p isogeny $\phi : A \to A'$ with $\mathrm{Ker}(\phi) = C$. We then consider $\underline{A}' = (A', \lambda', i', \eta'^{(p)})_{/S}$ given by $\lambda = \widehat{\phi} \circ \lambda' \circ \phi$, $i'(\alpha) = \phi \circ i(\alpha) \circ \phi^{-1}$, and $\eta'^{(p)} = \phi \circ \eta^{(p)}$. Then $\underline{A}'$ satisfies $\eta'^{(p)}(\widehat{L}^{(p)}) = \mathcal{T}^{(p)}(A)$.

If $\eta^{(p)}(\widehat{L}^{(p)}) \supset \mathcal{T}^{(p)}(A)$, we can find a prime-to-p isogeny $\underline{A}' \to \underline{A}$ such that $\phi \circ \eta'^{(p)} = \eta^{(p)}$, $\lambda' = \widehat{\phi} \circ \lambda \circ \phi$, $i'(\alpha) = \phi^{-1} \circ i(\alpha) \circ \phi$ and $\eta'^{(p)}(\widehat{L}^{(p)}) = \mathcal{T}^{(p)}(A)$. This fact follows from the canonical identification $\mathcal{T}(A) = \pi_1(A, \mathbf{0})$ for the origin $\mathbf{0}$ of A (see [ABV] Section 18).

If neither $\eta^{(p)}(\widehat{L}^{(p)}) \subset \mathcal{T}^{(p)}(A)$ nor $\eta^{(p)}(\widehat{L}^{(p)}) \supset \mathcal{T}^{(p)}(A)$, we can find two prime-to-p isogenies $\underline{A} \xrightarrow{\phi'} \underline{A}''$ and $\underline{A}' \xrightarrow{\phi} \underline{A}''$ for $\underline{A}'' = (A'', \lambda'', i'', \eta''^{(p)})_{/S}$ so that $\eta''^{(p)}(\widehat{L}^{(p)}) = \eta^{(p)}(\widehat{L}^{(p)}) \cap \mathcal{T}^{(p)}(A'')$ and $\eta'^{(p)}(\widehat{L}^{(p)}) = \mathcal{T}^{(p)}(A')$. Thus we can always find in the prime-to-p isogeny class of a given $\underline{A}_{/S}$, a quadruple $\underline{A}'_{/S}$ with $\eta'^{(p)}(\widehat{L}^{(p)}) = \mathcal{T}^{(p)}(A')$.

If $\phi : \underline{A}_{/S} \to \underline{A}'_{/S}$ is a prime-to-p isogeny with $\eta^{(p)}(\widehat{L}^{(p)}) = \mathcal{T}^{(p)}(A)$ and $\eta'^{(p)}(\widehat{L}^{(p)}) = \mathcal{T}^{(p)}(A')$, then $\deg(\phi) = 1$ and $\underline{A}_{/S} \cong \underline{A}'_{/S}$ by ϕ. Thus insisting $\eta^{(p)}(\widehat{L}^{(p)}) = \mathcal{T}^{(p)}(A)$, we can replace $\approx$ by $\cong$ in order to define the functor $\mathcal{E}_K^{(p)}$. In other words,

$$\mathcal{E}_K^{(p)}(S) \cong \mathcal{E}_K(S) = \left\{ \underline{A}_{/S} \middle| (\det) \text{ and } \eta^{(p)}(\widehat{L}^{(p)}) = \mathcal{T}^{(p)}(A) \right\} / \cong .$$

The functor defined in this way can be proven to be representable (as now show) by a $\mathcal{V}$-scheme $Sh_{K/\mathcal{V}}^{(p)}$.

Since $\widehat{\Gamma}^{(p\infty)}(N)$ for N prime to p gives a fundamental system of neighborhoods of the identity in $G(\mathbb{A}^{(p\infty)})$, we may assume that $K = \widehat{\Gamma}(N) =$

$\widehat{\Gamma}^{(p\infty)}(N) \times \widehat{\Gamma}_p$. We only need to show that $\mathcal{C}_{D,V}$ is relatively representable over the Mumford moduli $M_{\mathbb{Q}}$ given by Theorem 6.20. In order to emphasize the dependence on $(D, V, \langle \cdot, \cdot \rangle)$, we write $\mathcal{E}_K^D$ for the functor $\mathcal{E}_K$ with respect to K and D. Write $2d = \dim_{\mathbb{Q}}(V)$. Then $d = \dim_S A$ for $\underline{A}_{/S} \in \mathcal{E}_K^D(S)$, and d is independent of the choice of $\underline{A}_{/S}$ by (det). For simplicity, we assume that the polarization pairing $\langle\ ,\ \rangle$ in (L2) gives the self-duality of L. We then identify with $GSp_{2d}(\mathbb{Z})$ the similitude group of $\langle\ ,\ \rangle$ acting on L. In this way, we let $GSp_{2d}(R)$ act on $V(R) = V \otimes_{\mathbb{Q}} R$. Write K_0 for the open compact subgroup of $GSp_{2d}(\mathbb{A}^{(\infty)})$ maximal among those preserving $\widehat{L}$ and the level N structures (the group K_0 was written as $\widehat{\Gamma}(N)$ in Section 6.4.1). Then $K_0 \cap G(\mathbb{A}^{(\infty)}) = K$ and $K = \widehat{\Gamma}(N) \subset G(\mathbb{A}^{(\infty)})$. As described in Theorem 6.20, $\mathcal{E}_{K_0}$ is representable by a quasi-projective scheme $M = M_{\mathbb{Q}} = \mathfrak{M}_N$ defined over $\mathbb{Z}_{(p)}$. Let $\underline{\mathbf{A}} = (\mathbf{A}, \boldsymbol{\lambda}, \overline{\boldsymbol{\eta}}^{(p)})_{/M}$ be the universal quadruple over M with the universal abelian scheme $\mathbf{A} \in \underline{\mathbf{A}}$. We consider the functor from $M\text{-}SCH$ into $SETS$:

$$T_{/S} \mapsto \left[(\mathbf{A}_T, \boldsymbol{\lambda}_T, i : O_D \hookrightarrow \mathrm{End}_T(\mathbf{A}_T)) \middle| i(1_D) = \mathrm{id}_{\mathbf{A}},\ \boldsymbol{\lambda}_T^{-1} \circ {}^t i \circ \boldsymbol{\lambda}_T = i \circ *\right].$$

This functor is representable by a scheme $M_{D/M}$ (Corollary 6.11). Since the level structure $\overline{\boldsymbol{\eta}}^{(p)}$ on $\mathbf{A}$ gives rise to a level structure $\overline{\eta}^{(p)}$ of $(\mathbf{A}_T, i)$, we have a triple $\underline{A} = (\mathbf{A}_T, i, \overline{\eta}^{(p)})_{/T}$. Thus $\mathcal{E}_K^D$ is a subfunctor of the above functor. Again by the rigidity of endomorphisms under specialization (described in Section 6.1.6), $\mathcal{E}_K^D$ is represented by a closed subscheme $Sh_{K/\mathcal{V}}^{(p)}$ of M_D whose connected components are (each) finite over $M_{\mathbb{Q}}$. We show that $Sh_K^{(p)}$ is of finite type over $M_{\mathbb{Q}}$ (so it is projective and finite over $M_{\mathbb{Q}}$). Take a geometric point $x \in M_{\mathbb{Q}}$; suppose that we have $i : O_D \hookrightarrow \mathrm{End}_{\mathbf{A}}$ as above, which gives rise to a geometric point $y \in Sh_K^{(p)}$. For a given T, if T is connected, $\mathrm{End}(\mathbf{A}_T) \otimes_{\mathbb{Z}} \mathbb{Q}$ is a semi-simple algebra of finite dimension with positive involution (see [ABV] IV). Thus the number of embedding $D \hookrightarrow \mathrm{End}(\mathbf{A}_T) \otimes_{\mathbb{Z}} \mathbb{Q}$ is finite up to inner automorphisms. Moreover, the number is bounded by a constant only depending on the dimension of $\mathbf{A}$, that is d, because $\dim(\mathrm{End}_T(\mathbf{A}_T) \otimes_{\mathbb{Z}} \mathbb{Q})$ is bounded by $4d^2$ (e.g., [GME] Theorem 4.1.19). If one changes i by an inner automorphism induced by $\alpha \in \mathrm{End}(\mathbf{A}_T)$ and if we suppose that $(\mathbf{A}_T, \alpha i \alpha^{-1}, \lambda, \overline{\eta}^{(p)})$ is still an element of $\mathcal{E}_K^D(T)$, it is easy to see, by the condition that $\eta^{(p)}(\widehat{L}^{(p)}) = \mathcal{T}^{(p)}(A)$ combined with (L1-2), that α has to be an automorphism of $\mathbf{A}_T$. Since automorphisms of an abelian variety preserving a given polarization are only finitely many by the positivity of the polarization ([ABV] Theorem 5 in Section 21), there are only finitely many possibilities of having $i : O_D \hookrightarrow \mathrm{End}(\mathbf{A}_T)$ that give rise to an element of $\mathcal{E}_K^D(T)$. Thus $Sh_{K/\mathcal{V}}^{(p)} \to M_{\mathbb{Q}/\mathcal{V}}$ is quasi-finite. Then the projectivity of each connected component of $Sh_K^{(p)}$ over $M_{\mathbb{Q}}$ implies the finiteness of the map $Sh_K^{(p)} \to M_{\mathbb{Q}}$. If K is small, the morphism is unramified and finite. Actually, one can show that the morphism $Sh_K^{(p)} \to M_{\mathbb{Q}}$ is a closed immersion (over $\mathbb{Q}$) if N is sufficiently large (cf. [D1] 1.15; see Section 8.4.2 in the text).

The rigidity of endomorphism of abelian schemes under specialization combined with the deformation theory of abelian schemes tells us, for N sufficiently large, $Sh^{(p)}_{K/\mathcal{V}}$ is finite over $M_{\mathbb{Q}/\mathcal{V}}$ and smooth over $\mathcal{V}$ (see Theorem 8.8, the remark after the theorem, and [LaR] 6.2).

Although we assumed that L is self-dual, replacing GSp_{2d} by the similitude group of $\langle\cdot,\cdot\rangle$ (which is a conjugate in $GL(2d)$ of $GSp(2d)$) and using the Mumford moduli $M_{D,N}$ for the positive square root (Pffafian) D of the discriminant of $\langle\cdot,\cdot\rangle$, we can easily generalize the above argument to a given (nonself-dual) polarization of degree prime to p.

If D is a division algebra and $V = D$, the Shimura variety $Sh^{(p)}$ is known to be projective over $\mathcal{V}$. Indeed, if the image of M_D intersects with the cuspidal divisor of the toroidal compactification $\overline{M}_{\mathbb{Q}}$ described in Section 6.3.6, the pullback of the semi-abelian scheme $\mathcal{G}_{/\overline{M}_{\mathbb{Q}}}$ to M_D yields a non-abelian semi-abelian variety $\mathcal{G}_{D/k}$ with $D \subset \mathrm{End}^{\mathbb{Q}}(\mathcal{G}_{/k})$ as a fiber over a geometric point $Spec(k) \in M_D$. Since $\mathcal{G}_D$ is not an abelian scheme, its toric part $\mathcal{T} \subset \mathcal{G}_D$ is nontrivial and stable under all endomorphisms of $\mathcal{G}_D$ (rigidity of the toric part; see [NMD] 9.2). Thus D acts on its character group $X(T) \otimes_{\mathbb{Z}} \mathbb{Q}$, whose dimension over $\mathbb{Q}$ is less than $\dim \mathcal{G}_D = \frac{1}{2}\dim_{\mathbb{Q}} V$. This is impossible since $\mathrm{rank}_D V = 1$. Thus M_D never intersects with the cuspidal divisor of $\overline{M}_{\mathbb{Q}}$ and hence is projective (by the valuative criterion of properness). This finishes the proof of Theorem 7.1. We record here a byproduct of our proof.

Corollary 7.2 *Suppose the pairing in* (L1–2) *induces a self-duality on L. Then forgetting O_D-linear structure $\mathcal{E}^D_K \to \mathcal{E}^{\mathbb{Q}}_{K_0}$ induces a finite morphism of $Sh^{(p)}_K$ into $Sh^{(p)}_{K_0}(GSp(2d), \mathcal{X})$ and induces a finite unramified morphism if K is sufficiently small, where $2d = \dim_{\mathbb{Q}} V$ and $\mathcal{X}$ is the symmetric space of $GSp(2d)(\mathbb{R})$.*

We now replace prime-to-p isogenies in the definition of $\mathcal{E}^{(p)}_K$ by isogenies and prime-to-p structure $\eta^{(p)}$ by the full level structure $\eta : V(\mathbb{A}^{(\infty)}) \cong V(A)$. We define a functor $\mathcal{E}^{(\emptyset)}_K : E\text{-}SCH \to SETS$ given by

$$\mathcal{E}^{(\emptyset)}_K(S) = \left\{ \underline{A}_{/S} = (A, \overline{\lambda}, i, \overline{\eta})_{/S} \,\middle|\, A \text{ satisfies (det) and (pol) below} \right\} / \approx .$$

Here $\approx$ is up to isogeny (not up to prime-to-p-isogeny). Let $h_A : \mathbb{S} \to \mathrm{Aut}_{\mathbb{C}}(H_1(A, \mathbb{R}))$ for $\mathbb{S} = \mathrm{Res}_{\mathbb{C}/\mathbb{R}}\mathbb{G}_m$ be the homomorphism inducing the Hodge decomposition of $H_1(A, \mathbb{C})$ (see just above (pol) in Section 4.2.1). Note that $V_s(A) = H_1(A, \mathbb{A}^{(\infty)}) = H_1(A, \mathbb{Q}) \otimes_{\mathbb{Q}} \mathbb{A}^{(\infty)}$, and recall the condition (pol) adjusted to our setting:

(pol) There exists an D-linear isomorphism $f : V \cong H_1(A, \mathbb{Q})$ such that $f^{-1} \circ h_A \circ f \in X_G$, $(f \otimes 1_{\mathbb{A}^{(\infty)}}) \in (\eta \circ K)$ and $E_A(f(x), f(y)) = \alpha\langle x, y\rangle$ up to $\alpha \in \mathbb{Q}^{\times}$,

where E_A is the Riemann form on $H_1(A, \mathbb{Q})$ (see [ABV] Section 1). This condition really determines the group G as a similitude group of E_A under the identification by f.

In the same way, we can prove the representability of the version $\mathcal{E}_K^{(\emptyset)}$ of the functor $\mathcal{E}_K^{(p)}$ replacing in the definition of $\mathcal{E}_K^{(p)}$ prime-to-p-isogenies by isogenies. The new functor $\mathcal{E}_K^{(\emptyset)}$ therefore gives rise to the Shimura variety $Sh(G,X)$ defined over the reflex field E. Recall the warning we made after Theorem 7.1: the condition (pol) is not imposed in the definition of $\mathcal{E}_K^{(p)}$; so, $\mathcal{E}_K^{(p)}$ restricted to E-algebras may be a disjoint union of $\mathcal{E}_K^{(\emptyset)}$ for a different choice of an alternating form on V which is locally equivalent to $\langle\cdot,\cdot\rangle$. This subtle point is studied in Section 7.1.5. See Theorem 7.7 for the description of the variety $Sh(G,X)$.

Depending on the data at the cusps governing toroidal compactifications, a semi-abelian scheme $\mathcal{G}_{/\overline{M}_{\mathbb{Q}}}$ is constructed in [DAV] (see Section 6.3.6). Then even in the quasi-projective (but nonprojective) case, a similar argument using Corollary 6.11 (applied to $\mathcal{G}_{/\overline{M}_{\mathbb{Q}}}$ in place of $\mathbf{A}_{/M_{\mathbb{Q}}}$) gives a scheme projective over the smooth toroidal compactification $\overline{M}_{\mathbb{Q}}$ of the Mumford moduli $M_{\mathbb{Q}}$. Since the endomorphism algebra of an abelian variety $A_{/k}$ for an algebraically closed field k (after tensoring $\mathbb{Q}$) is semi-simple, there are only finitely many possibilities of embedding D into $\mathrm{End}_k(A)\otimes_{\mathbb{Z}}\mathbb{Q}$ up to conjugation. Thus the morphism $\overline{M}_D \to \overline{M}_{\mathbb{Q}}$ has finite geometric fiber everywhere; that is, the morphism is quasi-finite. Since the scheme $\overline{M}_D$ is proper over $\overline{M}_{\mathbb{Q}}$ (see Corollary 6.11), it has to be finite. Thus writing $\underline{\omega} = f_*\Omega_{\mathbf{A}/M_D}$ for $f : \mathbf{A} \to M_D$ and defining $\det(\underline{\omega})$ by its maximal exterior product, we can define a graded algebra:

$$\mathfrak{G}^K = \mathfrak{G}_D^K = \bigoplus_{n=0}^{\infty} H^0(\overline{M}_D, \det(\underline{\omega})^{\otimes n}).$$

Moreover, as seen in Section 6.3.6, $M_{\mathbb{Q}}^* = Proj(\mathfrak{G}_{\mathbb{Q}}^K)$ and hence we have the minimal compactification of Sh_K defined by $Sh_K^* = Proj(\mathfrak{G}_D^K)$ finite over $M_{\mathbb{Q}}^*$.

The smoothness of $\overline{M}_{D/\mathcal{V}}$ around the cusps for a well-chosen cuspidal datum was shown by Fujiwara for G and GU of types A and C ([F]). If we choose the cuspidal data for $GSp(2d)$ and G so that they are compatible (in other words, so that the pullback of the semi-abelian scheme over $\overline{M}_{\mathbb{Q}}$ is the semi-abelian scheme over $\overline{M}_D$ associated with the cuspidal data for G), this guarantees that the q-expansion parameter is well defined over $\mathcal{V}$ and projectivity of $\overline{M}_D$ over $\mathcal{V}$ (if the level is prime to p), because it is finite over projective $\overline{M}_{\mathbb{Q}}$.

Even if K is not very small, we always have a coarse moduli scheme $Sh_K^{(p)}$ representing the functor $\mathcal{E}_K^D$ over $\mathcal{V}$. The above arguments all work well. We write $\overline{Sh}_K$ for a toroidal compactification of Sh_K and Sh_K^* for the minimal compactification. Since the natural morphisms

$$Sh_K^* \to M_{\mathbb{Q}}^* \quad \text{and} \quad \overline{Sh}_K \to \overline{M}_{\mathbb{Q}}$$

are quasi-finite and projective, they are finite. Let V be the image of Sh_K^* in $M_{\mathbb{Q}}^*$. Then $V = Proj(\mathfrak{G}_K^*)$ for a graded algebra $\mathfrak{G}_K^*$ which is the quotient

of $\mathfrak{G}^K_{\mathbb{Q}}$. Then, assuming the existence of smooth toroidal compactification of Sh_K, we have, for any open compact subgroup $K \subset G(\mathbb{A}^{(\infty)})$ maximal at p,

$$Sh^*_K = Proj(\mathfrak{G}^K_D). \tag{Proj}$$

Here $\mathfrak{G}^K_D$ is the integral closure of $\mathfrak{G}^*_K$ in the total quotient ring of $\mathfrak{G}^*_K$. This follows from the fact that Sh^*_K is smooth outside cusps, and at the cusps, if K is sufficiently small, $\overline{M}_D$ is finite over $\overline{M}_{\mathbb{Q}}$ (and normal over $\mathcal{V}$). The graded algebra $\mathfrak{G}^K_D$ is the graded algebra of automorphic forms on G if $\dim Sh_K > 1$.

7.1.3 Moduli Variety for Similitude Groups

Recall that F_0 is the subfield of F fixed by the involution $*$. We write O_0 for the integer ring of F_0. We have formulated the moduli problem for the group G with similitude norm $\nu : G \to \mathbb{Q}$. But we can consider polarization λ modulo totally positive elements in $\mathcal{V}_0^\times$ for $\mathcal{V}_0 = O_0 \otimes_{\mathbb{Z}} \mathbb{Z}_{(p)}$. Then we obtain the classification problem for the F_0-similitude group $GU_{/\mathbb{Q}}$. Note that the connected component X^+ of the symmetric spaces of GU and G are identical and that GU has a nice property that its center $Z \cong \mathrm{Res}_{F/\mathbb{Q}}\mathbb{G}_m$ is cohomologically trivial (i.e., $H^1_{ct}(\mathrm{Gal}(M/F_0), Z(M)) = 0$ for all Galois extensions M/F_0; Hilbert's Theorem 90).

The group $\mathrm{Res}_{F/\mathbb{Q}}GL(2)$ (for $F = F_0$, $D = F$, and $V = F^2$) we considered in Section 4.2 is a group of this type; so, we have a detailed exposition of the Shimura variety of GU in this special case of Hilbert modular functions. Although we have written this group as GU, as this example shows, the group GU may not be a unitary group.

We now give some details of the Shimura variety for general GU. To define the moduli problem of abelian schemes associated with $GU_{/\mathbb{Q}}$, we need to modify slightly the morphisms of the fibered category $\mathcal{C}_{D,V}$. We define a new fibered category $\mathcal{A}^{(p)} = \mathcal{A}^{(p)}_{D,V}$ over $\mathcal{V}$-SCH to be the category of quadruples $\underline{A}_{/S} = (A, \lambda, i, \overline{\eta}^{(p)})$ for $\overline{\eta}^{(p)} = \eta^{(p)} \mod K$, where $K \subset GU(\mathbb{A}^{(\infty)})$ is a closed subgroup maximal at p. Recall that O_0 is the integer ring of F_0 and that $\mathcal{V}_0 = O_0 \otimes_{\mathbb{Z}} \mathbb{Z}_{(p)}$. Again $\mathcal{V}_0$ is a semi-local Dedekind domain under (unr). Then we define

$$\begin{aligned}&\mathrm{Hom}_{\mathcal{A}^{(p)}_{/S}}(\underline{A}_{/S}, \underline{A}'_{/S})\\ &\quad= \left\{ \phi \in \mathrm{Hom}(A, A') \otimes \mathbb{Z}_{(p)} \middle| \begin{array}{l} \widehat{\phi} \circ \lambda' \circ \phi = \lambda \circ i(a) \text{ with } a \in \mathcal{V}_{0+}^\times \\ \phi \circ i = i' \circ \phi \text{ and } \overline{\eta}'^{(p)} = \phi \circ \overline{\eta}^{(p)} \end{array} \right\}, \end{aligned} \tag{7.5}$$

where $\mathcal{V}_{0+}^\times$ indicates the group of totally positive units in $\mathcal{V}_0$. We then consider the functor

$$\mathcal{E}_K^{\mathcal{A}^{(p)}}(S) = \left\{ \underline{A}_{/S} \middle| A \text{ satisfies (det)} \right\} / \approx,$$

where $\approx$ indicates isomorphism classes in $\mathcal{A}^{(p)}_{D/S}$.

We may allow all isogenies as morphisms to define the category $\mathcal{A}^{(\emptyset)}$ (with full level structure) and define a functor from $E\text{-}SCH$ to $SETS$ by

$$\mathcal{E}_K^{\mathcal{A}^{(\emptyset)}}(S) = \left\{ \underline{A}_{/S} \middle| A \text{ satisfies (det) and (pol) below} \right\} / \approx,$$

where $\approx$ indicates isomorphism classes in $\mathcal{A}^{(\emptyset)}_{D/S}$. Note that $V_s(A)$ is canonically isomorphic to $H_1(A, \mathbb{A}^{(\infty)}) = H_1(A, \mathbb{Q}) \otimes_{\mathbb{Q}} \mathbb{A}^{(\infty)}$. Recall

(pol) There exists a D-linear isomorphism $f : V \cong H_1(A, \mathbb{Q})$ such that $(f \otimes 1_{\mathbb{A}^{(\infty)}}) \in (\eta \circ K)$, $f^{-1} \circ h_A \circ f \in X_G$, and $\langle f(x), f(y) \rangle_\lambda = \alpha \langle x, y \rangle_0$ up to $\alpha \in F_0^\times$,

where $\mathrm{Tr}_{F_0/\mathbb{Q}}(\langle \cdot, \cdot \rangle_\lambda) = E_A(\cdot, \cdot)$ for the Riemann form E_A on $H_1(A, \mathbb{Q})$ and similarly $\mathrm{Tr}_{F_0/\mathbb{Q}}(\langle \cdot, \cdot \rangle_0) = \langle \cdot, \cdot \rangle$.

In the Hilbert modular case, we constructed the modular variety first representing the functor $\mathcal{E}_{\Gamma(N)}$ with a fixed polarization. Since $\mathbb{Z}_+^\times = \{1\}$, fixing polarization is a moduli problem of $G(\mathbb{Q}) = \{g \in GL_2(F) | \det(g) \in \mathbb{Q}\}$, because of $\mathcal{E}^{(\emptyset)}_{\Gamma(N)} \cong \mathcal{E}_{\Gamma(N)}$ proven in Section 4.2.1. Then we extended the construction to $GU = \mathrm{Res}_{F/\mathbb{Q}} GL(2)$ by forming quotients of a disjoint union of the G-moduli $\mathfrak{M}_\mathfrak{c}$ by subgroups of $O_+^\times = \nu(GU(\mathbb{Z}))$ (like construction of $\mathfrak{M}_{\Gamma_1(\mathfrak{N})}$ in Section 4.1.2). In the general case of GU, the Hermitian symmetric domain X^+ is the same for G and GU. The group GU is the similitude group of the Hermitian pairing $\langle \cdot, \cdot \rangle_0 : V \times V \to F_0$ given by $\mathrm{Tr}_{F_0/\mathbb{Q}}(\langle \cdot, \cdot \rangle_0) = \langle \cdot, \cdot \rangle$. As we show in Section 7.1.5, the local equivalence class (the genus class) of the Hermitian alternating form $\langle \cdot, \cdot \rangle_0$ under GU is trivial, but modulo $G(\mathbb{Q})$, there may be several subclasses (giving rise to a finite number of unitary groups $G_{/\mathbb{Q}}$ with $G_i(\mathbb{A}) = G(\mathbb{A})$; see Remark 7.4). Following the same process (as in the Hilbert modular case), starting from a disjoint union $\bigsqcup_i Sh(G_i, X)$ of $Sh(G_i, X)$ (and making quotients by subgroups of $\nu(G_i(\mathbb{Z})) = \nu(G(\mathbb{Z}))$), the above functors $\mathcal{E}_K^{\mathcal{A}^{(\emptyset)}}$ and $\mathcal{E}_K^{\mathcal{A}^{(p)}}$ can be proved to be representable if K is sufficiently small. We write $Sh(GU, X)_{/E}$ (resp. $Sh^{(p)}(GU, X)_{/\mathcal{V}}$) with $X = X_{GU}$ for the scheme representing $\mathcal{E}_K^{\mathcal{A}^{(\emptyset)}}$ (resp $\mathcal{E}_K^{\mathcal{A}^{(p)}}$), because in this case, we prove that $Sh^{(p)} \times_{\mathcal{V}} E \cong Sh(GU, X)/GU(\mathbb{Z}_p)$ in Theorem 7.5.

The compactification of the moduli space M_K also can be done as described above. The only point we need to make explicit is that if the class $\overline{\lambda}$ of polarizations modulo multiplication of totally positive elements in $\mathcal{V}_0^\times$ is defined over S, we can always find a representative λ defined over S. In other words, for each geometric point $s \in S$, if the action of $\pi_1(S, s)$ preserves $\overline{\lambda}$, then the question is, "Can we find $\lambda : A \to {}^tA$ well defined over S?" Taking an étale finite Galois extension S'/S over which λ is well defined, we have a Galois 1-cocycle $c : \mathrm{Gal}(S'/S) \ni \sigma \mapsto \lambda^{1-\sigma}$ with values in the trivial $\mathrm{Gal}(S'/S)$-module $\mathcal{V}_{0+}^\times = \mathcal{V}_0^\times \cap F_{0+}^\times$. Since $\mathrm{Gal}(S'/S)$ is finite and $c : \mathrm{Gal}(S'/S) \to \mathcal{V}_{0+}^\times$ is a homomorphism with values in a torsion-free module, we find c is trivial, and λ is indeed defined over S.

We can also show that 2λ is globally induced by a line bundle on A relatively ample over S as follows: picking one symmetric polarization λ, the pullback $L^{\Delta}(\lambda)$ by $1\times\lambda$ of the universal line bundle over $A\times_S {}^tA/A$ (the Poincaré bundle) is always ample and is equal to 2λ (see [GIT] Proposition 6.10); so, in the class $\overline{\lambda}$, we can always find a polarization globally over S associated with a line bundle on A.

7.1.4 Classification of G

We define an algebraic group G_{0/F_0} by $G_0(R)=\left\{x\in C\otimes_{F_0}R\middle| xx^\iota=1\right\}$ for F_0-algebras R. Then we have $U=\mathrm{Res}_{F_0/\mathbb{Q}}G_0$. The involution $*$ of D either induces a nontrivial involution on F (a positive involution of the second kind) or the identity map on F (a positive involution of the first kind). If $*$ is of the second kind, F is a CM field over the totally real F_0, and $*$ coincides on F with the unique nontrivial automorphism over F_0 (see [Sh] and [ABV] Section 21). Then G_0 is an inner form of a quasi-split unitary group over F_0. We call this case "Case A" and call the group G "type A" (following the classification of Lie algebras; see [BLI]).

When $*$ induces the identity map on F, then $F=F_0$ is totally real, and the group G_0 is an inner form of either the symplectic group (Case C and the group of type C) or an orthogonal group of even variables (Case D and the group of type D). We have (cf. [Sh] Sections 1 and 2 or [Ko] Section 4)

$$C_\infty\cong\begin{cases} M_k(\mathbb{C})^{I_0} \text{ and } x^\iota=I_{s,t}{}^t\overline{x}I_{s,t} & \text{in Case A,}\\ M_{2k}(\mathbb{R})^{I_0} \text{ and } x^\iota=J_k{}^txJ_k^{-1} & \text{in Case C,}\\ M_k(\mathbb{H})^{I_0} \text{ and } x^\iota=-i^t\overline{x}i & \text{in Case D,}\end{cases}\tag{7.6}$$

where $\mathbb{H}=\mathbb{R}+\mathbb{R}i+\mathbb{R}j+\mathbb{R}k$ is the algebra of Hamilton quaternions, I_0 is the set of all field embeddings of F_0 into $\mathbb{R}$, $J_k=\left(\begin{smallmatrix}0 & -1_k\\ 1_k & 0\end{smallmatrix}\right)$ and $I_{s,t}=\left(\begin{smallmatrix}1_s & 0\\ 0 & -1_t\end{smallmatrix}\right)$ are as before, and $x\mapsto\overline{x}$ is either complex conjugation or quaternion conjugation.

There are type-B (orthogonal) groups admitting Shimura varieties, but they are non-PEL type except for the case when the signature is either $(2,2)$ or $(2,1)$ (and hence the groups are isogenous to the product of inner forms of $PGL(2)$). In Case D, the simple connectedness (sc) fails, but the argument proving Theorem 7.1 still (almost) works assuming $p>2$ (see [Ko]), and we get $Sh^{(p)}_{/\mathcal{V}}$. When $K^{(p)}$ is sufficiently small, Sh_K is smooth over $\mathcal{V}$. This follows from the fact that the deformation ring of a quadruple $(A,\lambda,i,\overline{\eta}^{(p)})$ is always formally smooth (see [LaR] 6.2, also Theorem 8.8 in the text and the remark after the theorem). If $C=\mathrm{End}_D(V)$ is a division algebra ($\Leftrightarrow$ D is a division algebra and $\mathrm{rank}_D V=1$), as we have already seen, $Sh^{(p)}_{K/\mathcal{V}}$ is projective.

Since the stabilizer $C_{\mathbf{0}}$ of the fixed $h_{\mathbf{0}}\in X$ in $G(\mathbb{R})$ is a maximal compact subgroup modulo center by (pos), the connected component of X is isomorphic to the symmetric domain $\mathcal{H}=G(\mathbb{R})/C_{\mathbf{0}}$. An explicit form of $\mathcal{H}$ as a classical bounded matrix domain is given in [Sh] (see also [ACM] Chapter VI for the

domains in Cases A and C), along with an explicit method of constructing all possible analytic families of abelian varieties over the domain. We have computed $\mathcal{H}$ already in Section 5.2.2 for unitary and symplectic groups (which are groups of types A and C). The complex analytic space $Sh_K(\mathbb{C})$ is given by $G(\mathbb{Q})\backslash G(\mathbb{A})/KC_{\mathbf{0}} = G(\mathbb{Q})\backslash(X \times G(\mathbb{A}^{(\infty)}))/K$, and its connected component is given by $\Gamma\backslash\mathcal{H}$ for $\Gamma = (gKg^{-1}G(\mathbb{R})^+) \cap G(\mathbb{Q})$ with a suitable $g \in G(\mathbb{A}^{(\infty)})$.

7.1.5 Generic Fiber of $Sh_K^{(p)}$

In the definition of the functor $\mathcal{E}_K^{(p)}$, the structure of the Tate module $V^{(p)}(A) = H_1(A, \mathbb{Q}) \otimes_{\mathbb{Q}} \mathbb{A}^{(p\infty)}$ is specified modulo K for a test object $(A, \overline{\lambda}, \eta^{(p)})$ over a characteristic 0 field inside $\mathbb{C}$. However, this information may not uniquely determine the D-Hermitian structure of $H_1(A, \mathbb{Q})$ which supplies us with the exact information of the algebraic group giving rise to the Shimura variety; so, the variety $Sh_K^{(p)}$ representing the functor $\mathcal{E}_K^{(p)}$ may be a disjoint union of the reduction modulo p of characteristic 0 Shimura varieties associated with different groups locally isomorphic to each other (but not globally). We study this point in this section to relate $Sh^{(p)}$ to the characteristic 0 fiber of the Shimura variety.

A $(D, *)$-Hermitian form $(\cdot, \cdot)$ is a symmetric bilinear form on V with values in F_0 such that $(bx, y) = (x, b^*y)$ for all $x, y \in V$. The given alternating form $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{Q}$ can be uniquely lifted to an F_0-linear alternating form $\langle \cdot, \cdot \rangle_0 : V \times V \to F_0$ so that $\langle \cdot, \cdot \rangle = \mathrm{Tr}_{F_0/\mathbb{Q}}(\langle \cdot, \cdot \rangle_0)$. Pick $(A, \overline{\lambda}, \overline{\eta}^{(p)})_{/k} \in \mathcal{E}_K^{\mathcal{A}^{(p)}}(k)$ for an extension field k/E inside $\mathbb{C}$. Similarly, the Riemann form E_A on $H_1(A(\mathbb{C}), \mathbb{Q})$ with values in $\mathbb{Q}$ can be uniquely lifted to an F_0-linear alternating form $\langle \cdot, \cdot \rangle_\lambda$ so that $\mathrm{Tr}_{F_0/\mathbb{Q}}(\langle \cdot, \cdot \rangle_\lambda) = E_A(\cdot, \cdot)$. Then the test object $(A, \overline{\lambda}, \overline{\eta}^{(p)})_{/k}$ supplies us with the following information.

> The D-module $V_A := H_1(A(\mathbb{C}), \mathbb{Q})$ with F_0-linear Riemann form $\langle \cdot, \cdot \rangle_\lambda : V_A \times V_A \to F_0$ and an isomorphism $\eta^{(p)} : V \otimes_{\mathbb{Q}} \mathbb{A}^{(p\infty)} \cong V_A \otimes_{\mathbb{Q}} \mathbb{A}^{(p\infty)}$ sending $\langle \cdot, \cdot \rangle_0$ to $\langle \cdot, \cdot \rangle_\lambda$ up to the constant in $O^\times_{0,(p)}$.

The Shimura variety $Sh(GU, X)_{/E}$ is defined for the group GU; so, our question is if the above data are sufficient to determine GU (i.e., to verify (pol) using information only from defining data of $\mathcal{E}_K^{\mathcal{A}^{(p)}}$). There are three points to check:

(1) $(V_{\mathbb{R}}, \langle \cdot, \cdot \rangle_0) \cong (V_A \otimes_{\mathbb{Q}} \mathbb{R}, \langle \cdot, \cdot \rangle_\lambda)$, and for any two ι-homomorphisms $\mathbb{C} \hookrightarrow C_\infty$ are conjugates under $G(\mathbb{R})$;
(2) $(V_p, \langle \cdot, \cdot \rangle_0) \cong (V_A \otimes_{\mathbb{Q}} \mathbb{Q}_p, \langle \cdot, \cdot \rangle_\lambda)$;
(3) the Hasse principle for the alternating form $\langle \cdot, \cdot \rangle_0$.

First we prove (1) following [Ko], Section 2. Recall that I_0 is the set of all real embeddings of F_0 (regarded as the set of all Archimedean places of F_0). For any semi-simple F-algebra B and a B-module M, we put $B_\sigma = B \otimes_{F_0,\sigma} \mathbb{R}$ and $M_\sigma = M \otimes_{F_0,\sigma} \mathbb{R}$; so, $B_{\mathbb{R}} = B \otimes_{\mathbb{Q}} \mathbb{R} = \prod_{\sigma \in I_0} B_\sigma$ and $M_{\mathbb{R}} = \prod_{\sigma \in I_0} M_\sigma$.

Each D_σ has positive involution induced by $*$, which we still denote by $*$. Now a $(D,*)$-Hermitian form on V_σ is a symmetric $\mathbb{R}$-bilinear form with $(bv,w)=(v,b^*w)$ for all $v,w\in V_\sigma$. We start with

Lemma 7.3 *Let $(\cdot,\cdot)$ and $(\cdot,\cdot)'$ be two positive-definite $(D,*)$-Hermitian forms on V_σ. Then there exists $c\in C_\sigma^\times$ such that $(cv,cw)=(v,w)'$ for all $v,w\in V_\sigma$.*

Proof. The Hermitian form $(\cdot,\cdot)$ (resp. $(\cdot,\cdot)'$) induces an identification $\phi: V_\sigma\cong \widehat{V}_\sigma:=\mathrm{Hom}_{\mathbb{R}}(V_\sigma,\mathbb{R})$ (resp. $\phi': V_\sigma\cong\widehat{V}_\sigma$). We regard $\widehat{V}_\sigma$ as a D_σ-module by $f\mapsto f\circ b^*$. Since $\phi(bv)(w)=(bv,w)=(v,b^*w)=\phi(b^*w)$, we find $c_1=\phi^{-1}\circ\phi'$ is an element in $\mathrm{Aut}_{D_\sigma}(V_\sigma)=C_\sigma^\times$. We find $(c_1v,w)=\phi(cv)(w)=\phi(\phi^{-1}\circ\phi'(v))(w)=(v,w)'$. Since the two forms are $(D,*)$-Hermitian, we find $c_1=c_1^\iota$. Consider the set $C_+=\{c\in C_\sigma | c^\iota=c \text{ and } c>0\}$, where $c>0\Leftrightarrow(cv,w)$ is positive-definite. Any positive-definite symmetric inner products of a finite-dimensional real vector space are equivalent over $\mathbb{R}$ (Sylvester's theorem), C_+ is therefore isomorphic to a positive subcone of the space of positive-definite symmetric matrices of $\mathrm{End}_{\mathbb{R}}(V_\sigma)$. In particular, $\log: C_+\cong\log(C_+)$ and $\exp:\log(C_+)\cong C_+$ are well defined by the standard power series expansions. Therefore, we have $c=\sqrt{c_1}=\exp(\log(c_1/2))$ in C_+, which gives the desired element satisfying $(cv,cw)=(c^2v,w)=(c_1v,w)=(v,w)'$. □

If we have two ι-homomorphism $h,h':\mathbb{C}\hookrightarrow C_\sigma$, we have two Hermitian forms $(v,w)=\langle v,h(i)w\rangle_\sigma$ and $(v,w)'=\langle v,h'(i)w\rangle_\sigma$. We find $c\in C^\times$ with $\langle v,h'(i)w\rangle_\sigma=(v,w)'=(cv,cw)=\langle cv,h(i)cw\rangle_\sigma=\langle v,c^\iota h(i)cw\rangle_\sigma$. This shows that $h'(z)=c^\iota h(z)c$. Since $h(i)h(-i)=h(i)h(i)^\iota=1$, we verify $cc^\iota=1$ because $cc^\iota\in C_+$; so, h and h' are conjugates under $G_0(F_{0,\sigma})$.

The Hodge decomposition $H^1(A,\mathbb{C})=H^0(A,\Omega_{A/\mathbb{C}})\oplus H^0(A,\overline{\Omega}_{A/\mathbb{C}})$ induces, by Poincare duality, an isomorphism $h_A:\mathbb{C}\to\mathrm{End}(C_{\mathbb{R}})$ so that $h_A(z)$ acts on $H(A,\Omega_{A/\mathbb{C}})$ by multiplication by z and on $H(A,\overline{\Omega}_{A/\mathbb{C}})$ by multiplication by $\overline{z}$. By the theory of Lefschetz–Poincaré ([ABV] Section 1), the Hermitian pairing $\langle x,h_A(i)y\rangle_\lambda$ is positive-definite. Since $V_A\otimes_{\mathbb{Q}}\mathbb{Q}_\ell\cong V\otimes_{\mathbb{Q}}\mathbb{Q}_\ell$ for a prime ℓ, $V\cong V_A$ as D-modules by faithfully flat descent for $\mathbb{Q}_\ell/\mathbb{Q}$. Thus applying the above argument to $(v,w)=\langle v,h(i)w\rangle$ for $h\in X$ and $(v,w)'=\langle v,h_A(i)w\rangle_\lambda$, we conclude h_A is a conjugate of h under $G_1(\mathbb{R})$; so, $h_A\in X$. Now we apply the above argument to (v,w) as above and $(v,w)'=\langle v,h(i)w\rangle_\lambda$; we conclude (1).

As for (3), we consider a functor $\mathcal{G}:k\mapsto\mathrm{Isom}((V\otimes_{\mathbb{Q}}k,\langle\cdot,\cdot\rangle),(V_A,\langle\cdot,\cdot\rangle_\lambda))$ for number fields k. Here the isomorphism preserves the alternating forms up to elements in $(\mathrm{Res}_{F_0/\mathbb{Q}}\mathbb{G}_m)(k)=(k\otimes_{\mathbb{Q}}F_0)^\times$. The group GU acts on the functor by $\phi\mapsto\phi\circ g$; so, obviously $\mathcal{G}$ gives a GU-torsor (i.e., a principal homogeneous space over GU) defined over $\mathbb{Q}$. As is well known (see [GCH] I.5 and III.1 and [Sp]), the GU-torsors are classified by the (non-abelian) Galois cohomology set $H^1(\mathbb{Q},GU)=H^1(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}),GU(\overline{\mathbb{Q}}))$. Thus the proof of (3) boils down to showing injectivity of the (diagonal) restriction map

$H^q(\mathbb{Q}, GU) \to H^q(\mathbb{R}, GU) \times \prod_\ell H^q(\mathbb{Q}_\ell, GU)$ for $q = 1$ (the Hasse principle for $H^q(\mathbb{Q}, GU)$).

For groups of types B and D, G_1 is not simply connected; so, often the Hasse principle fails (cf. [GCH] III.4.7 and [Kn1]). Thus we assume one of the following two conditions which guarantees the condition (sc).

(B1) The involution $*$ on D is a positive involution of the second kind (Case A);
(B2) The involution $*$ is of the first kind, and $D_{\mathbb{R}} = D \otimes_{\mathbb{Q}} \mathbb{R} \cong M_d(F_{\mathbb{R}})$ for $F_{\mathbb{R}} = F \otimes_{\mathbb{Q}} \mathbb{R}$ (Case C).

These conditions eliminate groups of types B and D.

We now prove (3) under (B1) or (B2). We consider the exact sequence: $1 \to G_1 \to GU \to Z^{ZU} \to 1$ for cocenter Z^{GU} of GU. Then by the long exact sequence associated with the above short exact sequence, we get the following exact sequence (see [Sp] Proposition 2.2), $H^1(\mathbb{Q}, G_1) \to H^1(\mathbb{Q}, GU) \to H^1(\mathbb{Q}, Z^{GU})$. Since G_1 is simply connected (of type A or C), it is known that $H^1(\mathbb{Q}, G_1)$ is trivial (see [Kn1]). Thus we need to show either that $H^1(\mathbb{Q}, Z^{GU}) = 0$ or the Hasse principle holds for $H^1(\mathbb{Q}, Z^{GU})$. We write $T_K = \mathrm{Res}_{K/\mathbb{Q}}\mathbb{G}_m$ for a number field K. When the group is of type C, $Z^{GU} \cong T_F$, because GU is either $\mathrm{Res}_{F/\mathbb{Q}}GSp(2g)$ for totally real F or a similitude group of a Hermitian form over a totally indefinite quaternion algebra over F (see [We]). Then $H^1(\mathbb{Q}, Z^{GU}) = 0$ for type C groups GU follows from Hilbert's Theorem 90.

If the group GU is of type A, it is a unitary group; so, taking r with $C_\sigma \cong M_r(\mathbb{C})$, we find $G_1 = \{x \in C | N_C(x) = \nu(x) = 1\}$ for the similitude norm $\nu(x) = xx^\iota$ and the reduced norm $N_C : C \to F$. Since C is of dimension r^2 over F, we find that $N_C(x)N_C(x)^c = \nu(x)^r$ for the generator c of $\mathrm{Gal}(F/F_0)$. This shows that $Z^{GU} = \{(x, y) \in T_F \times T_{F_0} | N_{F/F_0}(x) = y^r\}$. If $r = 2m$ is even, we have $Z^{GU} \cong T_1 \times T_{F_0}$ by sending $(x, y) \mapsto (xy^{-m}, y) \in T_1 \times T_{F_0}$ for $T_1 = \mathrm{Ker}(N_{F/F_0} : T_F \to T_{F_0})$. From the exact sequence, $1 \to T_1 \to T_F \to T_{F_0} \to 1$, we have another $1 \to \mathrm{Ker}(N_{F/F_0}) \to F^\times \xrightarrow{N_{F/F_0}} F_0^\times \to H^1(\mathbb{Q}, T_1) \to 1$. The vanishing of $H^1(\mathbb{Q}, T_F)$ follows from Hilbert's Theorem 90. This shows that $H^1(\mathbb{Q}, Z^{GU}) \cong H^1(\mathbb{Q}, T_1) \cong F_0^\times / N_{F/F_0}(F^\times)$ for which the Hasse principle holds (the norm theorem of class field theory for cyclic extensions). Thus the Hasse principle holds for $H^1(\mathbb{Q}, GU)$ as desired.

Now assume that r is odd. Let $Z_1 \cong \mu_r$ be the center of G_1. Then we have an exact sequence of groups $1 \to Z_1(\overline{\mathbb{Q}}) \to Z(\overline{\mathbb{Q}}) \to Z_{GU}(\overline{\mathbb{Q}}) \to 1$, because over $\overline{\mathbb{Q}}$, $GU \cong (GL(r) \times \mathbb{G}_m)^{I_0}$. Since $H^2(\mathbb{Q}, \mu_r)$ is the kernel of multiplication by r of the Brauer group of $\mathbb{Q}$ (Kummer's theory), the Hasse principle holds for $H^2(\mathbb{Q}, \mu_r)$ (class field theory; see [BNT] XIII.3). For the same reason, Hasse's principle holds for $H^2(\mathbb{Q}, T_F)$. Since $Z = T_F$, by Hilbert's Theorem 90, $H^1(\mathbb{Q}, Z) = 0$. Then by the short exact sequence, $0 = H^1(\mathbb{Q}, Z) \to H^1(\mathbb{Q}, Z^{GU}) \to H^2(\mathbb{Q}, \mu_r) \to H^2(\mathbb{Q}, Z)$, we conclude the Hasse principle for $H^1(\mathbb{Q}, Z^{GU})$ and hence for $H^1(\mathbb{Q}, GU)$. This shows (3) for

GU. The Hasse principle sometimes fails for G of type A (but in a rather harmless way; see Remark 7.4 and [Ko] Section 7).

Now we prove (2). Let $O_C = \{x \in C | xL = L\}$. We need to prove that if L'_p is a self-dual $O_{D,p}$-lattice of $V_p = V \otimes_{\mathbb{Q}} \mathbb{Q}_p$ under $\langle \cdot, \cdot \rangle_0$, then we have an element $g \in GU(\mathbb{Q}_p)$ such that $gL_p = L'_p$. We can do this at each place $\mathfrak{p}$ of F_0 above p. By (unr), $O_{D,\mathfrak{p}} = \prod_{\mathfrak{P}|\mathfrak{p}} M_d(O_{\mathfrak{P}})$. This shows that $O_{C,\mathfrak{p}} = \prod_{\mathfrak{P}|\mathfrak{p}} M_r(O_{\mathfrak{P}})$. Thus we may assume that $d = 1$ (replacing L by $\varepsilon L \cong O_{\mathfrak{p}}^d$ for the $*$-invariant idempotent $\varepsilon = \mathrm{diag}[1, 0, \ldots, 0] \in O_{D,\mathfrak{p}}$). Then the existence of $g \in U(\mathbb{Q}_p) \subset GU(\mathbb{Q}_p)$ as above follows from [EPE], Lemma 5.9 (note here any self-dual lattice of $V_{\mathfrak{p}}$ is maximal in the sense of [EPE], Chapter I). There is another proof in [Ko], Section 7 reducing the result to a theorem of S. Lang on the vanishing of $H^1(\mathbb{F}_q, H)$ for a connected algebraic group H over a finite field $\mathbb{F}_q$. When the group G is of type D, G may not be connected; so, the proof of (2) for such groups may also fail.

Since the proof in the case of groups of type C is easy, let us give more details of Case C. In this case, $L_{\mathfrak{p}} = O_{\mathfrak{p}}^r$ for an even integer r. Choose $h \in C$ so that $h'L_p = L'_p$. Then the $(x, y) \mapsto \langle h'x, h'y \rangle_0$ induces a new skew Hermitian form $\langle \cdot, \cdot \rangle_1$ on L_p; so, we need to find $h \in O_C^\times$ such that $\langle hx, hy \rangle_1 = \langle x, y \rangle_0$. Take $x_1, y_1 \in L_{\mathfrak{p}}$ with $\langle x_1, y_1 \rangle_0$ a unit u in $O_{\mathfrak{p}}$. Changing y_1 by $u^{-1}y_1$, we may assume that $\langle x_1, y_1 \rangle_0 = 1$. Then $L_{\mathfrak{p}} = (O_{\mathfrak{p}}x_1 + O_{\mathfrak{p}}y_1) \oplus (O_{\mathfrak{p}}x_1 + O_{\mathfrak{p}}y_1)^\perp$. By induction on $\mathrm{rank}\, L_{\mathfrak{p}}$, we find that $\langle \cdot, \cdot \rangle_0$ is equivalent to $(x, y) \mapsto xJ_{r/2}{}^t y$ on $O_{\mathfrak{p}}^r$. Similarly, $\langle \cdot, \cdot \rangle_1$ is equivalent to $(x, y) \mapsto xJ_{r/2}{}^t y$ on $O_{\mathfrak{p}}^r$, and we find the desired $h \in O_C$ with $\langle hx, hy \rangle_1 = \langle x, y \rangle_0$.

Remark 7.4 If G is of type A, Kottwitz proved that $Sh_K^{(p)} \otimes E$ is a union of the Shimura varieties $Sh(G_i, X)/K_p$ for different unitary groups G_i with $G_i(\mathbb{A}^{(\infty)}) \cong G(\mathbb{A}^{(\infty)})$ indexed by $i \in \mathrm{Ker}(H^1(\mathbb{Q}, G) \to \prod_\ell H^1(\mathbb{Q}_\ell, G))$. Thus if Hasse's principle fails for $H^1(\mathbb{Q}, G)$, the functor $\mathcal{E}_K^{(p)}$ is bigger than $\mathcal{E}_K^{(\emptyset)}$ over E-algebras (see [Ko], Section 8), although $Sh(G, X)$ and $Sh(GU, X)$ share the same neutral component. Thus to have the identity $Sh^{(p)} \otimes_{\mathcal{V}} E = Sh/G(\mathbb{Z}_p)$, we need to have the cohomological triviality condition ((CT) in the following subsection) for the center Z of the group G.

We have proved that $Sh^{(p)}(GU, X) \otimes_{\mathcal{V}} E = Sh(GU, X)/GU(\mathbb{Z}_p)$ for GU of types A and C, because we recover the polarization condition (pol) from the data defining $\mathcal{E}_K^{\mathcal{A}^{(p)}}$. For each $(A_0, \overline{\lambda}_0, \eta_0^{(p)})$ over $\overline{\mathbb{F}}_p$, by deformation theory of abelian schemes due to Grothendieck–Messing ([LaR] 6.2, [Zi], and [CBT]; see also Theorem 8.8 and the remark after that), we can lift the triple $(A_0, \overline{\lambda}_0, \eta_0^{(p)})$ to a triple $(A, \overline{\lambda}, \eta^{(p)})$ over $W = W(\overline{\mathbb{F}}_p)$ (the ring of Witt vectors with coefficients in $\overline{\mathbb{F}}_p$). Indeed, there is a canonical lift if A_0 is ordinary corresponding to the Serre–Tate coordinate $q = 0$ under (unr) (see Sections 8.2.4 and 8.2.5). Then for the quotient field $\mathcal{K}$ of W, we find that $(A, \overline{\lambda}, \eta^{(p)}) \in \mathcal{E}^{(\emptyset)}_{GU(\mathbb{Z}_p)}(\mathcal{K})$, and the universal deformation space of $(A_0, \overline{\lambda}_0, \eta_0^{(p)})$ is a formal spectrum of

a power series ring over W, and hence $Sh^{(p)}(GU, X)$ is smooth over $\mathcal{V}$. We have proven

Theorem 7.5 *Suppose* (unr) *and one of the conditions* (B1–2). *Then the p-integral model* $Sh^{(p)}(GU, X)$ *representing* $\mathcal{E}_1^{\mathcal{A}^{(p)}}$ *is smooth over* $O_{E,(p)}$, *and we have* $Sh^{(p)}(GU, X) \otimes_{O_{E,(p)}} E = Sh(GU, X)/GU(\mathbb{Z}_p)_{/E}$, *where* $GU(\mathbb{Z}_p) = \{x \in GU(\mathbb{Q}_p) | xL_p = L_p\}$.

7.2 General Shimura Varieties

We describe canonical models of Shimura varieties of PEL and non-PEL types as well as the global and local reciprocity laws characterizing the variety in a succinct way for our later use.

7.2.1 Axioms Defining Shimura Varieties

We first formulate Shimura varieties as in Deligne [D1] and [D2], and in many places, our exposition follows Milne's expository article [Mi], Chapter II. We consider a reductive group $G_{/\mathbb{Q}}$ with $G(\mathbb{R})^+/C_{\mathbf{0}}$ isomorphic to a Hermitian symmetric domain X^+. We write the adjoint group of G as $G^{ad} = G/Z$ in this section (which was also written as PG before). We can identify X^+ with a connected component of the set X of the $G(\mathbb{R})$-conjugacy class of a fixed homomorphism (of real algebraic groups) $h_{\mathbf{0}} : \mathbb{S}_{/\mathbb{R}} \hookrightarrow G_{/\mathbb{R}}$ whose stabilizer (under conjugation) is $C_{\mathbf{0}}$. Each maximal compact subgroup of $G_1(\mathbb{R})$ can be given as the centralizer of such a homomorphism. In other words, we have the origin $\mathbf{0} \in X^+$ and the corresponding homomorphism $h_{\mathbf{0}} : \mathbb{S} \hookrightarrow G$ with centralizer $C_{\mathbf{0}}$ in $G(\mathbb{R})$. We write $x \in X$ when we consider x geometrically as a point of the symmetric domain of G and write $h_x : \mathbb{S} \hookrightarrow G_{/\mathbb{R}}$ when we consider $x \in X$ as a homomorphism.

We start with such a pair (G, X) and require (G, X) to satisfy the following four axioms.

(D1) For each $x \in X$ the Hodge structure on $Lie(G)$ defined by the homomorphism h_x is of type $\{(-1,1),(0,0),(1,-1)\}$. In other words, $Ad(h_x(z))$ $(z \in \mathbb{C}^\times = \mathbb{S}(\mathbb{R}))$ has only eigenvalues $z^{-1}\overline{z}$, 1, and $z\overline{z}^{-1}$.

(D2) For each $x \in X$, $g \mapsto h_x(i)gh_x(i)^{-1}$ is a Cartan involution of the adjoint group $G^{ad}_{/\mathbb{R}} = PG_{/\mathbb{R}}$; so, the centralizer of h_x in $G_1(\mathbb{R})$ is a maximal compact subgroup. In other words, the involution gives the twisted complex conjugation (which we referred to in the introduction) making the centralizer of h_x the subgroup of real points of a real algebraic group.

(D3) The adjoint group $G^{ad}_{/\mathbb{Q}}$ has no factor defined over $\mathbb{Q}$ whose real points form a compact group (this condition excludes a case we treated in Section 4.3: $G = G_D = \mathrm{Res}_{F/\mathbb{Q}} D^\times$ for a definite quaternion algebra D over a totally real field F).

(D4) The identity component $Z^\circ = Z^\circ(G)$ of the center $Z = Z(G)$ of G splits over a CM field.

In addition to these conditions, we assume the following two extra conditions to avoid technicality.

(SC) The derived group G_1 of G is simply connected (this condition excludes some orthogonal groups of type B and groups of type D but includes their Spin and Clifford covering groups).
(CT) The center Z of G is a cohomologically trivial torus.

Here (CT) means the vanishing of the Galois cohomology $H^1(F, Z) = H^1_{ct}(\mathrm{Gal}(\overline{\mathbb{Q}}/F), Z(\overline{\mathbb{Q}}))$ for each finite extension $F/\mathbb{Q}$ (see [MFG] Chapter 4 for Galois cohomology). Thus the group GU satisfies the condition (CT) as we already remarked but the group G in (7.1) may not.

Example 7.6 We quote the following three types of the pairs (G, X) satisfying the above six conditions (D1–4), (SC) and (CT) from [MiS].

A. (Unitary groups: type A groups) Let $(D_{/F}, *)$ be a central simple algebra over a totally imaginary quadratic extension F over a totally real field F_0 (a CM field over F_0) having a positive involution $*$ inducing nontrivial automorphism c on F/F_0. We write I_0 for the set of real embedding of the totally real field F_0. The group $G_{/\mathbb{Q}}$ is given by

$$G(k) = \left\{ g \in (D \otimes_{\mathbb{Q}} k)^\times \middle| xx^\iota \in (F_0 \otimes_{\mathbb{Q}} k)^\times \right\}$$

for $\mathbb{Q}$-algebras k. Then $G(\mathbb{R}) \cong \prod_{\sigma \in I_0} GU(m_\sigma, n_\sigma)$, where

$$GU(m, n) = \left\{ x \in GL_{m+n}(\mathbb{C}) \middle| x I_{m,n} {}^t\overline{x} = \nu(g) I_{m,n} \right\}$$

with $I_{m,n} = \left(\begin{smallmatrix} 1_m & 0 \\ 0 & -1_k \end{smallmatrix}\right)$. Then X is the $G(\mathbb{R})$ conjugacy class of $h_{\mathbf{0}} : \mathbb{S} = \mathbb{C}^\times \hookrightarrow G(\mathbb{R})$ given by $z \mapsto \left(\left(\begin{smallmatrix} z 1_{m_\sigma} & 0 \\ 0 & \overline{z} 1_{n_\sigma} \end{smallmatrix} \right) \right)_\sigma$. These are included in the groups we have defined in (7.1) (this is the case where $V = D$ and hence $C = {}^t D$ under the notation of (7.1)).

B. (Orthogonal groups: type B groups) Let q be a quadratic form on a vector space V over a totally real field F such that it has signature

$$(m_\sigma, n_\sigma) \in \{(d, 0), (0, d), (d-2, 2), (2, d-2)\}$$

on $V_\sigma = V \otimes_{F,\sigma} \mathbb{R}$ ($d = \dim_F V$). We suppose that $\dim_F V \geq 3$. We write I_s for the places of F where q has signature $(d-2, 2)$ or $(2, d-2)$. We put $I_r = I - I_s$. Then $G = \mathrm{Res}_{F/\mathbb{Q}} Gpin(q)$ for the special Clifford group $Gpin(q)_{/F}$ for q. By definition,

$$G_1(\mathbb{R}) \cong \prod_{\sigma \in I_s} Spin(d-2, 2)(\mathbb{R}) \times \prod_{\sigma \in I_r} Spin(d, 0)(\mathbb{R}).$$

We follow [Si] for the description of the origin $h_{\mathbf{0}} : \mathbb{S} \hookrightarrow G(\mathbb{R})$ of X. Decompose $h_{\mathbf{0}} = \prod_\sigma h_\sigma$ and $h_\sigma = 1$ if $\sigma \in I_r$. Write

$$\nu : G = \mathrm{Res}_{F/\mathbb{Q}} Gpin(q) \twoheadrightarrow \mathrm{Res}_{F/\mathbb{Q}}(Gpin(q)/Spin(q)) = \mathrm{Res}_{F/\mathbb{Q}}\mathbb{G}_{m/F}$$

for the lift of the similitude map by $Gpin(q) \twoheadrightarrow GO(q)$, and define the involution $\iota \circlearrowright G$ by $gg^\iota = \nu(g)$. For $\sigma \in I_s$, take two orthogonal vectors $e_1, e_2 \in V \otimes_{F,\sigma} \mathbb{R}$ with $q(e_j) = -1$ for $j = 1, 2$. Then in the Clifford algebra, $j_\sigma = e_1 e_2$ satisfies $j_\sigma^2 = -1$ and $j_\sigma^\iota = -j_\sigma$. Then $h_\sigma(a + b\sqrt{-1}) = a + bj_\sigma \in Gpin(q_\sigma)$, where q_σ is the quadratic form induced on $V \otimes_{F,\sigma} \mathbb{R}$ from q.

C. (Symplectic groups: type C groups) Let $D_{/F}$ be a quaternion algebra (we allow $D = M_2(F)$) for a totally real field F. Let V be a free left D-module of rank d. Write ι for the involution on D given by $x + x^\iota = \mathrm{Tr}(x)$ and $xx^\iota = N(x)$ for the reduced norm and trace $N, \mathrm{Tr} : D \to F$. Let $I_D = \{\sigma \in I | D \otimes_{F,\sigma} \mathbb{R} \cong M_2(\mathbb{R})\}$ and $I^D = I - I_D$. We suppose that $I_D \neq \emptyset$. We consider a nondegenerate symmetric form $(\ ,\ ) : V \times V \to F$ such that $(bx, y) = (x, b^\iota y)$. Then, for $\mathbb{Q}$-algebras k,

$$G(k) = \left\{ g \in \mathrm{Aut}_{D \otimes_{\mathbb{Q}} k}(V \otimes_{\mathbb{Q}} k) \middle| (gx, gy) = \nu(g)(x, y) \right\},$$

where $\nu(g) \in (F \otimes_{\mathbb{Q}} k)^\times$ for $g \in G(k)$; that is, $\nu : G \to \mathrm{Res}_{F/\mathbb{Q}}\mathbb{G}_m$ is a morphism of group schemes. We assume that $(\ ,\ )$ induces a positive- or negative-definite symmetric form on $V \otimes_{F,\sigma} \mathbb{R}$ if $\sigma \in I^D$. Then $G(\mathbb{R}) \cong \prod_{\sigma \in I_D} GSp_{2d}(\mathbb{R}) \times \prod_{\sigma \in I^D} GU_{\mathbb{H}}(d, 0)$, and the space X is the $G(\mathbb{R})$-conjugacy class of $h_{\mathbf{0}} : \mathbb{C}^\times \ni z = a + b\sqrt{-1} \mapsto \left(\begin{pmatrix} a1_d & b1_d \\ -b1_d & a1_d \end{pmatrix} \right)_{\sigma \in I_D} \times (1)_{\sigma \in I^D}$. Here

$$GU_{\mathbb{H}}(d, 0) = \{x \in GL_d(\mathbb{H}) | xx^\iota = \nu(x) \in \mathbb{R}^\times\}$$

for the algebra $\mathbb{H}$ of Hamilton quaternions. If there is no component isomorphic to $GU_{\mathbb{H}}$ in $G(\mathbb{R})$, this type of group is included in the groups we have defined in (7.1).

We fix an origin $\mathbf{0} \in X^+$, which we assume to be the one given in the above examples when we are in Cases A, B, and C. We first consider the Shimura variety over $\mathbb{C}$ in Deligne's sense,

$$M_{\mathbb{C}}(G, X)(\mathbb{C}) = \varprojlim_K G(\mathbb{Q}) \backslash G(\mathbb{A}) / K C_{\mathbf{0}}, \tag{7.7}$$

where $C_{\mathbf{0}}$ is the centralizer of $h_{\mathbf{0}}$ in $G(\mathbb{R})$ and K runs through open-compact subgroups of $G(\mathbb{A}^{(\infty)})$. This projective limit can be identified set-theoretically with

$$M_{\mathbb{C}}(G, X)(\mathbb{C}) = G(\mathbb{Q}) \backslash \left(X \times G(\mathbb{A}^{(\infty)}) \right) / \overline{Z(\mathbb{Q})}, \tag{7.8}$$

where the action is given by $\gamma(x, g)z = (\gamma x, \gamma x z)$ ($\gamma \in G(\mathbb{Q})$ and $z \in \overline{Z(\mathbb{Q})}$) and $\overline{Z(\mathbb{Q})}$ is the closure of $Z(\mathbb{Q})$ in $G(\mathbb{A}^{(\infty)})$ (see [D2] Proposition 2.1.10 and

[Mi] pages 323–324). We write $[x, g]$ for the complex point of the Shimura variety represented by the pair $x \in X$ and $g \in G(\mathbb{A}^{(\infty)})$. When G is given by GU of type A or C in (7.1), $GU(\mathbb{Q})/GU(\mathbb{Q})_+$ acts on $\pi_0(X)$ transitively; so, we can rewrite (7.8) as

$$M_{\mathbb{C}}(G, X)(\mathbb{C}) = G(\mathbb{Q})_+ \backslash \left(X^+ \times G(\mathbb{A}^{(\infty)}) \right) / \overline{Z(\mathbb{Q})}.$$

We recall the amalgamated product $G(\mathbb{A}^{(\infty)}) *_{G(\mathbb{Q})_+} G^{ad}(\mathbb{Q})^+$ defined in (4.30). We let $G^{ad}(\mathbb{Q})$ act on $G(\mathbb{A}^{(\infty)})$ by conjugation, and we make a semi-direct product $G(\mathbb{A}^{(\infty)}) \rtimes G^{ad}(\mathbb{Q})^+$, where $G^{ad}(\mathbb{Q})^+ = G^{ad}(\mathbb{Q}) \cap G^{ad}(\mathbb{R})^+$ for the identity-connected component $G^{ad}(\mathbb{R})^+$ of $G^{ad}(\mathbb{R})$. We write $G(\mathbb{Q})_+$ for the subgroup of $G(\mathbb{Q})$ whose image in $G^{ad}(\mathbb{Q})$ falls in the identity-connected component of $G^{ad}(\mathbb{R})$. We embed $G(\mathbb{Q})$ into the semi-direct product $G(\mathbb{A}^{(\infty)}) \rtimes G^{ad}(\mathbb{Q})$ by $\gamma \mapsto (\gamma^{-1} \rtimes ad(\gamma))$. Then the image of $G(\mathbb{Q})$ is a normal subgroup of $G(\mathbb{A}^{(\infty)}) \rtimes G^{ad}(\mathbb{Q})^+$. We then define $G(\mathbb{A}^{(\infty)}) * G^{ad}(\mathbb{Q})^+$ by $G(\mathbb{A}^{(\infty)}) \rtimes G^{ad}(\mathbb{Q})^+ / G(\mathbb{Q})$. This group acts on $M_{\mathbb{C}}(G, X)$ on the right by

$$[x, g](g' \rtimes \gamma) = [\gamma^{-1}(x), (gg')^{\gamma}],$$

where $g^t = t^{-1}gt$. The action is well defined because we have

$$[\gamma^{-1}\alpha(x), (\alpha g g')^{\gamma}] = [\alpha^{\gamma}\gamma^{-1}(x), \alpha^{\gamma}(gg')^{\gamma}] = [\gamma^{-1}(x), (gg')^{\gamma}] = [x, g](g' \rtimes \gamma)$$
$$\text{and} \quad [x, g](\gamma^{-1} \rtimes ad(\gamma)) = [\gamma^{-1}(x), (g\gamma^{-1})^{\gamma}] = [\gamma^{-1}(x), \gamma^{-1}g]$$

for $\alpha, \gamma \in G(\mathbb{Q})$.

7.2.2 Reciprocity Law at Special Points

Since $\mathbb{S}_{/\mathbb{C}} \cong \mathbb{G}_m \times \mathbb{G}_m$ by $z \mapsto (z, \overline{z})$, for any morphism of algebraic groups $\rho : G \to GL(V)$ defined over $\mathbb{Q}$ for a $\mathbb{Q}$-vector space V, the composition $\rho \circ h_x$ ($x \in X$) gives a double grading on $V \otimes_{\mathbb{Q}} \mathbb{C}$ (a Hodge structure). Deligne's viewpoint is to regard the Shimura variety over $\mathbb{C}$ to be a modulus of such Hodge structures with additional restrictions ([D2] Section 1). More boldly, one might expect that a given Shimura variety $Sh(G, X)$ is a fine modulus of a specific family of motives, since each motive has a natural Hodge structure of its Hodge realization (see [D2] Introduction, [Mi] II.3, and [Mi2]). In the case of a Shimura variety of PEL type, by construction, it is a fine modulus of specific abelian schemes up to isogeny; so, it is a modulus of the motives $H^1(A)$ associated with abelian varieties A. Via this heuristic, we expect that $M_{\mathbb{C}}(G, X)$ has a canonical model defined over a specific number field, the reflex field we now introduce.

We can pull back h_x to $\mathbb{G}_m$ by putting $\mu_x(z) = h_x(z, 1)$, getting a homomorphism $\mu_x : \mathbb{G}_m \hookrightarrow G$ defined over $\mathbb{C}$. Let $\mathcal{M}_X(\mathbb{C})$ be the conjugacy class of the homomorphism $\mu_{\mathbf{0}} : \mathbb{G}_m \hookrightarrow G$. Then $\mathcal{M}_X(\mathbb{C})$ is the set of complex points of an algebraic subvariety of G. Let $E = E(G, X)$ be the smallest field of definition of $\mathcal{M}_X$, that is, the fixed field of

$$\{\sigma \in \mathrm{Aut}(\mathbb{C}) | \sigma(\mathcal{M}_X(\mathbb{C})) = \mathcal{M}_X(\mathbb{C})\}.$$

Thus E is a finite extension of $\mathbb{Q}$, which is known to be a totally real field or a CM field. This follows from the classification of the pairs (G, X) satisfying (D1–3) done in [D2] 1.3.8, showing that $E(G, X)$ is a subfield of the fixed field K of the subgroup of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ fixing the Dynkin diagram D associated with the pair (G, X) ([D2] 2.3.5). In $\mathrm{Aut}(D)$, complex conjugation is in its center, and hence K is either a totally real field or a CM field ([IAT] 5.11). The field E is called the *reflex field* of (G, X) (which used to be called the dual field).

A point $x \in X$ is called a special point if there exists a maximal torus $T_x \subset G$ defined over $\mathbb{Q}$ such that h_x has values in $T_x(\mathbb{R})$. A point $[x, g] \in M_{\mathbb{C}}(G, X)$ is called *special* if $x \in X$ is special. By definition, we then have $\mu_x : \mathbb{G}_m \to T_x$. Let E_x be the minimal field of definition of μ_x, and write E_x^{ab} for the maximal abelian extension of E_x. The field E_x is a CM field containing E given by the fixed field of $\{\sigma \in \mathrm{Aut}(\mathbb{C}) | \sigma \circ \mu_x = \mu_x \circ \sigma\}$. This induces the following homomorphism of algebraic groups,

$$N_x : \mathrm{Res}_{E_x/\mathbb{Q}}\mathbb{G}_m \xrightarrow{\mathrm{Res}_{E_x/\mathbb{Q}}(\mu_x)} \mathrm{Res}_{E_x/\mathbb{Q}}T_x \xrightarrow{Norm} T_x \hookrightarrow G.$$

We then define the reciprocity map $r_x : \mathrm{Gal}(E_x^{ab}/E_x) \to T_x(\mathbb{A}^{(\infty)})/\overline{T_x(\mathbb{Q})}$ as follows. For each $\tau \in \mathrm{Gal}(E_x^{ab}/E_x)$, take an idele $t \in (E_x \otimes_{\mathbb{Q}} \mathbb{A})^\times$ giving rise to τ^{-1} by the Artin reciprocity map. Then define $r_x(\tau) = (N_x(t^{(\infty)}) \bmod \overline{T_x(\mathbb{Q})})$. It is easy to check that the image only depends on τ.

A scheme $Sh(G, X)$ over $E = E(G, X)$ is called a *canonical model* or the *Shimura variety* of $M_{\mathbb{C}}(G, X)$ if the following two conditions are met:

(rc1) $$Sh(G, X) \otimes_E \mathbb{C} \cong M_{\mathbb{C}}(G, X);$$

(rc2) If $[x, g] \in M_{\mathbb{C}}(G, X)$ is special, then $[x, g] \in Sh(G, X)(E_x^{ab})$ under the above isomorphism and $\tau([x, g]) = [x, r_x(\tau)g] = [x, g \cdot r_x(\tau)^g]$ for all $\tau \in \mathrm{Gal}(E_x^{ab}/E_x)$, where $r_x(\tau)^g = g^{-1}r_x(\tau)g$.

The condition (rc2) for $G = \mathrm{Res}_{F/\mathbb{Q}}GL(2)$ (for a totally real field $F = F_0$) is the local reciprocity law in Section 4.2.3, but in this general case, it is a requirement for $Sh(G, X)_{/E}$ to be the Shimura variety of (G, X). When (G, X) comes from the algebra $D_{/F}$ and D-module V as in Section 7.1, special points are those points on $Sh(G, X)$ carrying abelian varieties of CM-type, and (rc2) follows from the main theorem of complex multiplication (Theorem 4.19) as we have proven (rc2) (i.e., Corollary 4.20) for the Hilbert modular variety.

By (7.8), we find $[x, \gamma g] = [\gamma^{-1}x, g]$ ($\gamma \in G(\mathbb{Q})$). For $\gamma \in T_x(\mathbb{Q})$, $\gamma(x) = x$ if x is special with respect to T_x. Thus $[x, \gamma r_x(\tau)a] = [\gamma^{-1}(x), r_x(\tau)a] = [x, r_x(\tau)a]$ is well defined without ambiguity. Since the set of special points is Zariski dense (actually, it is dense also in complex topology; see [D1] 5.1). In all three cases (A, B, C), the existence of the canonical models is known (see [Sh5], [Si], [D1] 4.21, 6.4, and [D2] 2.7.20). More generally, if the pair (G, X) has a symplectic embedding into $(GSp(2N), \mathfrak{H}_N \sqcup \overline{\mathfrak{H}}_N)$ as in [D2] 2.3.2-3, the complex proanalytic space $M_{\mathbb{C}}(G, X)$ admits a unique canonical model

$Sh(G,X)_{/E(G,X)}$ (see [D2] 2.7.10). This type of Shimura variety is called a Shimura variety *of abelian type*. See [Mi] II for a discussion covering a more general type of Shimura variety.

7.2.3 Shimura's Reciprocity Law

In this subsection, we reformulate Shimura's tower of geometrically connected canonical models in Deligne's language. Each member satisfies an intricate relation under the action of the group $G(\mathbb{A}^{(\infty)})$ as described in [Sh5] II.5 and [MiS] Theorem 4.6, but we only describe the tower; so, we do not discuss this point here.

We now define, following [D2], $\rho : \widetilde{H} \to H_1$ for the universal covering $\widetilde{H}$ of the derived group H_1 of each reductive group $H_{/\mathbb{Q}}$. Then we put

$$\pi(H) = H(\mathbb{A})/H(\mathbb{Q})\rho(\widetilde{H}(\mathbb{A}))$$

and write $\pi_0\pi(H)$ for the group of connected components of $\pi(H)$. We put $\overline{\pi}_0\pi(H) = \pi_0\pi(H)/\pi_0(H(\mathbb{R})_+)$, where $H(\mathbb{R})_+$ is the inverse image of the identity component $H^{ad}(\mathbb{R})^+$ of $H^{ad}(\mathbb{R})$. We consider the following exact sequence

$$1 \longrightarrow G_1 \longrightarrow G \xrightarrow{\chi} Z^G \longrightarrow 1 \tag{7.9}$$

for a torus Z^G (called the *cocenter* of G) defined over $\mathbb{Q}$. Recall $\mu_x = h_x|_{\mathbb{G}_m \times 1} : \mathbb{G}_{m/\mathbb{C}} \to G_{/\mathbb{C}}$. For each $x \in X$, we look at the composite $\mu_X = \chi \circ \mu_x : \mathbb{G}_m \to Z^G$. Since $\{\mu_x\}_{x\in X}$ belongs to one conjugacy class X, μ_X is independent of x and is defined over $E = E(G,X)$ by the definition of the reflex field E. We can thus define $\lambda : \mathrm{Res}_{E/\mathbb{Q}}\mathbb{G}_m \to Z^G$ by the composite of $\mathrm{Res}_{E/\mathbb{Q}}\mathbb{G}_m \xrightarrow{\mathrm{Res}(\mu_X)} \mathrm{Res}_{E/\mathbb{Q}}Z^G$ and $\mathrm{Res}_{E/\mathbb{Q}}Z^G \xrightarrow{norm} Z^G$. We define a homomorphism

$$\ell : \pi_0\pi(\mathrm{Res}_{E/\mathbb{Q}}\mathbb{G}_m) = \pi_0(E_{\mathbb{A}}^\times/E^\times) \cong \mathrm{Gal}(E^{ab}/E) \to \overline{\pi}_0\pi(G)$$

following [MiS], page 929. Here we identify $\pi_0(E_{\mathbb{A}}^\times/E^\times)$ with $\mathrm{Gal}(E^{ab}/E)$ sending a prime element $\varpi \in F^\times \subset F_{\mathfrak{p}}^\times$ at a prime ideal $\mathfrak{p}$ to the geometric Frobenius at $\mathfrak{p}$ (thus compatible with the Artin reciprocity law; see Theorem 2.12). This homomorphism was defined by Deligne in [D2] (2.6.2.1) as $\pi_0 N_{E/\mathbb{Q}} q_M$ (using an intricate but explicit argument in terms of torsors without assuming (SC), [D2] 2.4–6). By (SC), the strong approximation theorem affirms (see [D1] Theorem 2.4) that $\pi_0\pi(G) \cong \pi_0\pi(Z^G) = \pi_0(Z^G(\mathbb{A})/Z^G(\mathbb{Q}))$ via χ. The homomorphism λ gives rise to a homomorphism of $\pi_0\pi(\mathrm{Res}_{E/\mathbb{Q}}\mathbb{G}_m) = \pi_0(E_{\mathbb{A}}^\times/E^\times)$ into $\pi_0\pi(Z^G) = \pi_0(Z^G(\mathbb{A})/Z^G(\mathbb{Q}))$. Thus we have

$$\overline{\pi}_0\pi(G) = Z^G(\mathbb{A})/\overline{Z^G(\mathbb{R})^+ Z^G(\mathbb{Q})}\chi(Z(\mathbb{R})).$$

Composing λ with the projection: $\pi_0\pi(Z^G) = \pi_0\pi(G) \twoheadrightarrow \overline{\pi}_0\pi(G)$, we obtain

$$\ell : \pi_0\pi(\mathrm{Res}_{E/\mathbb{Q}}\mathbb{G}_m) = \pi_0(E_{\mathbb{A}}^{\times}/E^{\times}) \xrightarrow{\lambda} \pi_0\pi(Z^G) = \pi_0\left(\frac{Z^G(\mathbb{A})}{Z^G(\mathbb{Q})}\right) \twoheadrightarrow \overline{\pi}_0\pi(G).$$

Let $\mathfrak{k}$ be the subfield of E^{ab} fixed by $\mathrm{Ker}(\ell)$. We then define a topological subgroup $\overline{\mathcal{G}}_+ = \mathcal{G}(G,X) \subset G(\mathbb{A})_+$ by the commutativity of the following diagram with two exact rows,

$$\begin{array}{ccccccc} 1 & \longrightarrow & \overline{G(\mathbb{Q})_+}G(\mathbb{R})_+ & \longrightarrow & G(\mathbb{A})_+ & \longrightarrow & \overline{\pi}_0\pi(G) \longrightarrow 1 \\ & & \| \uparrow & & \uparrow & & \cup \uparrow \ell \\ 1 & \longrightarrow & \overline{G(\mathbb{Q})_+}G(\mathbb{R})_+ & \longrightarrow & \mathcal{G}(G,X) & \xrightarrow[\sigma]{} & \mathrm{Gal}(\mathfrak{k}/E) \longrightarrow 1. \end{array} \tag{7.10}$$

Here $\overline{G(\mathbb{Q})_+}$ is the closure in $G(\mathbb{A})$ of $G(\mathbb{Q})_+ = G(\mathbb{Q}) \cap (G(\mathbb{A}^{(\infty)})G(\mathbb{R})_+)$, and ℓ restricted to $\mathrm{Gal}(\mathbb{Q}^{cyc}E/E)$ coincides with (the inverse of) the Artin reciprocity map (cf. Theorem 2.12). In Shimura's paper [Sh5], II, the notation $\overline{\mathcal{G}}_+$ is used for the group $\mathcal{G}(G,X)$, which is the notation used in [MiS], 4.13. As is shown in [MiS], 4.12–13, $\mathcal{G}(G,X)$ gives the stabilizer in $G(\mathbb{A})_+$ of the geometrically irreducible component of $Sh(G,X)$ containing the image of the identity of $G(\mathbb{A})$.

Let $\mathcal{Z}$ be the set of subgroups K of $\mathcal{G}(G,X)$ satisfying the following two conditions,

1. $\overline{Z(\mathbb{Q})G(\mathbb{R}_+)} \subset K$;
2. $K/\overline{Z(\mathbb{Q})G(\mathbb{R})_+}$ is open in $\mathcal{G}(G,X)$.

We recall that we fixed an origin $\mathbf{0} \in X^+ \subset X$. We then identify X with $G(\mathbb{R})/C_{\mathbf{0}}$ for the centralizer $C_{\mathbf{0}}$ of $h_{\mathbf{0}}$. The origin is taken to be $h_{\mathbf{0}}$ as exhibited in (A–C) in Example 7.6. We take the geometrically connected component $V_{/\mathfrak{k}}$ of $Sh(G,X)$ containing the image of the identity of $G(\mathbb{A})$ and define $V_{K/k_K} = V/K$ as a geometrically irreducible variety defined over k_K, where k_K is the fixed field of $\sigma(K)$ in $\mathfrak{k}$. Then this family of geometrically irreducible varieties $\{V_K\}_{K\in\mathcal{Z}}$ satisfies many fine properties as described in [Sh5] II.5 (see also [MiS] Theorem 4.6), which characterize the tower $\{V_K\}_{K\in\mathcal{Z}}$ of canonical models.

Define $\mathfrak{K} = \bigcup_{K\in\mathcal{Z}} k_K(V_K)$ for the function field $k_K(V_K)$ of Sh_K over k_K. Then we have the following global and local reciprocity law (we have proved this at least for Hilbert modular varieties in Theorem 4.14 and Siegel modular varieties in Theorem 6.26).

Theorem 7.7 (Shimura, Deligne, Milne–Shih) *There exists a canonical group homomorphism $\tau : \mathcal{G}(G,X) \to \mathrm{Aut}(\mathfrak{K}/E)$ characterized by the following five properties.*

(1) *The kernel of τ is $\overline{Z(\mathbb{Q})}G(\mathbb{R})_+$;*
(2) *For $\alpha \in \mathcal{G}(G,X)$, $\tau(\alpha)|_{\mathfrak{k}} = \sigma(\alpha)$;*
(3) *For any $\alpha \in G(\mathbb{Q})_+ \subset \mathcal{G}(G,X)$ and $z \in Sh(G,X)$, $\tau(\alpha)(f(z)) = f(\alpha^{-1}(z))$ for all $f \in \mathfrak{K}$;*

(4) *For each special point* $z \in X^+$ *and* $f \in \mathfrak{K}$, $f(z) \in E_z^{ab}$ *if* f *is defined at* z, *where* X^+ *is the connected component of* X *containing* $\mathbf{0}$. *Moreover, if* f *is defined at special* $z \in X_+$, *then* $\tau(r_z(\rho))(f)$ *is defined at* z *and* $\rho(f(z)) = \tau(r_z(\rho))(f(z))$;

(5) *The map* τ *defines a topological isomorphism of* $\mathcal{G}(G,X)/\overline{Z(\mathbb{Q})}G(\mathbb{R})_+$ *onto an open subgroup of* $\mathrm{Aut}(\mathfrak{K}/E)$. *For every* $K \in \mathcal{Z}$, *we have* $\tau(K) = \mathrm{Aut}(\mathfrak{K}/k_K(Sh_K))$. *Here* $\mathrm{Aut}(\mathfrak{K}/E)$ *is equipped with the Krull topology as described in* Section 2.3.

This theorem has been proven by Shimura in [Sh5], II in Case C. After some work by students of Shimura, the general case in the theorem is covered by Milne–Shih [MiS], Theorem 4.13. Our formulation here follows [MiS] and is different from Shimura's formulation in which he uses $\tau(x^{-1})$ for τ in the present notation to make Galois action as a right action. We followed Shimura's formulation in Theorem 2.43, and this point is obvious from $\tau(\alpha)(f) = f \circ \alpha$ in Theorem 2.43 which is the inverse of $\tau(\alpha)$ in (3) above. Note here that the group $\tau(\mathcal{G}(G,X))$ is not just an open subgroup of the field automorphisms of $\mathfrak{K}$ but also the subgroup of scheme automorphisms of $Sh(G,X)$ over E.

The cokernel of τ also can be described explicitly as a subgroup of $\mathrm{Aut}(F/\mathbb{Q})$ in Cases A, B, and C, and in these cases, one can extend the group $\mathcal{G}(G,X)$ to describe explicitly the full automorphism group $\mathrm{Aut}(\mathfrak{K}/E)$ (see [Sh5] II, [MiS] 4.7 and 4.14, and [H03b]). In the proof of [MiS], one of the key ingredients of the description of the full group is the determination of $\mathrm{Aut}(\mathbb{C}\mathfrak{K}/\mathbb{C})$ by Miyake [Mt], and in [H03b], the theorem of Faltings and Zarhin (Theorem 4.13) is used in place of Miyake's result.

8

Ordinary p-Adic Automorphic Forms

We first describe a formal theory of false automorphic forms and find a set of conditions for nearly ordinary p-adic cusp forms to have vertical control. After this, we describe deformation theory by Serre–Tate of abelian varieties in order to prove the divisibility by the character value $\mu(\xi)$ in (5.40) of the Hecke operator associated with an element ξ of the expanding semi-group D in Section 5.1.2. By this, we have well-defined p-integral Hecke operators $T_p(\xi)$ on coherent cohomology groups and therefore well-defined nearly p-ordinary automorphic forms. Verifying the set of conditions on sections of automorphic vector bundles of a given (projective) Shimura variety of PEL type, we prove the vertical control theorem for automorphic forms on the Shimura varieties. We also state a similar result for quasi-projective Shimura varieties of quasi-split unitary groups without a detailed proof. At the end, we prove the irreducibility of the Igusa tower over the ordinary locus of symplectic and unitary Shimura varieties, reducing the problem to the Siegel modular case. By means of the Serre–Tate deformation coordinates, even when the Shimura variety does not have a cusp, we are able to define an analogue of the p-adic valuation v (in (4.36)) of its arithmetic automorphic function field; so, the irreducibility can be proven even for projective symplectic and unitary Shimura varieties of PEL type.

8.1 True and False Automorphic Forms

We describe the theory of false automorphic forms for a formal analogue of the Igusa tower. In our application, we remove nonordinary locus from the modulus $M_{/W_m}$ of abelian schemes of PEL type and write S_{m/W_m} for the ordinary locus. The scheme for $0 < m < \infty$ is obtained by inverting the Hasse invariant. For $m = \infty$, we define a formal scheme $S_{\infty/W}$ by $S_\infty = \varprojlim_m S_m$. If we choose a lift $E \in H^0(M_{/W}, \underline{\omega}^a)$ of the Hasse invariant and define $S_{/W} = M[\frac{1}{E}]_{/W}$, S_∞ is the formal completion of S along its special fiber modulo p. In this setting, sections in $H^0(S_\infty, \underline{\omega}^\kappa)$ are called *false* automorphic forms. On the

other hand, sections in $H^0(M, \underline{\omega}^\kappa)$ are called *true* or *"classical"* automorphic forms. The theory is a formal generalization of the work of Deligne–Katz in the elliptic modular case (see [K1] Appendix III).

The main purpose of this section is threefold:

1. Approximate p-adic automorphic forms by finite sums of classical forms;
2. Define the p-ordinary projectors e and e_0;
3. Find a set of (axiomatic) conditions that guarantees the VCT.

8.1.1 An Axiomatic Igusa Tower

Recall the mixed characteristic complete discrete valuation ring W with residue characteristic p. Let ϖ be a uniformizing parameter of W. Set $W_m = W/p^m W$. Let S be a flat W-scheme. We put $S_m = S \times_W W_m$. Then S_m is a sequence of flat W_m-schemes, given with isomorphisms: $S_{m+1} \otimes_{W_{m+1}} W_m \cong S_m$. We write S_∞ for the formal completion $\varprojlim_m S_m$ of S. Let P be a rank g p-adic étale sheaf on the S_ms; thus, $P_{/S_{m+1}}$ induces $P_{/S_m}$, $P = \varprojlim_n P/p^n P$, and $P_n = P/p^n P$ is an étale twist of the constant sheaf $(\mathbb{Z}/p^n\mathbb{Z})^g$.

We can slightly generalize our setting and could suppose that there exists a finite extension $F/\mathbb{Q}$ with integer ring $O = O_F$ and a homomorphism $O \hookrightarrow \mathrm{End}_{S_m}(P)$ such that $P_n \cong (O/p^n O)^g$ for all n locally under étale topology. Since we can transfer any of our results to this slightly general situation, just replacing $\mathbb{Z}_p$ by $O_p = O \otimes_{\mathbb{Z}} \mathbb{Z}_p$, we state our result only for P with $P_n \cong (\mathbb{Z}/p^n\mathbb{Z})^g$. This simplification also allows us to save some symbols.

We write $\pi_{m,n} : T_{m,n} = \mathrm{Isom}_{S_m}(P_n, (\mathbb{Z}/p^n\mathbb{Z})^g) \to S_m$ for the finite étale S_m-scheme representing the following functor, $SCH_{/S_m} \to SETS$,

$$(\pi : X \to S_m) \mapsto \left\{ \text{isomorphisms } \psi_n : P_{n/X} \cong (\mathbb{Z}/p^n\mathbb{Z})^g_{/X} \right\}.$$

The representability follows from the theory of Hilbert schemes as we have seen (see Section 6.1.5). Indeed, by $\psi \mapsto \psi^{-1}$, we have

$$\mathrm{Isom}_{S_m}(P_n, (\mathbb{Z}/p^n\mathbb{Z})^g) \cong \mathrm{Isom}_{S_m}((\mathbb{Z}/p^n\mathbb{Z})^g, P_n).$$

Since $\psi : (\mathbb{Z}/p^n\mathbb{Z})^g \cong P_n$ is determined by $(\psi(e_1), \ldots, \psi(e_n))$ for the standard base $(e_1, \ldots, e_n)$ of $(\mathbb{Z}/p^n\mathbb{Z})^g$, we find that $\mathrm{Isom}_{S_m}((\mathbb{Z}/p^n\mathbb{Z})^g, P_n)$ is represented by the open subscheme of P_n^g given by

$$\left\{ (\psi(e_1), \ldots, \psi(e_n)) \in P_n^g \;\middle|\; \psi(e_1) \wedge \cdots \wedge \psi(e_n) \neq 0 \right\}. \tag{8.1}$$

Because P_n^g / S_m is étale, $T_{m,n/S_m}$ is étale. Since each geometric fiber of $T_{m,n}$ over S_m is isomorphic to $GL_g(\mathbb{Z}/p^r\mathbb{Z})$ everywhere, it is faithfully flat and finite. Therefore $T_{m,n}$ is affine over S_m. We define $V_{m,n} = H^0(T_{m,n}, \mathcal{O}_{T_{m,n}})$.

The group $GL_g(\mathbb{Z}/p^n\mathbb{Z})$ acts on $T_{m,n}$ by $\psi \mapsto x\psi$ for $x \in GL_g(\mathbb{Z}/p^n\mathbb{Z})$ without fixed points, and we have $T_{m,n}/\Gamma_{n,n'} \cong T_{m,n'}$ for all $n' \leq n$, where

$$\Gamma_{n,n'} = \left\{ x \in GL_g(\mathbb{Z}/p^n\mathbb{Z}) \middle| x \equiv 1 \mod p^{n'} \right\}.$$

Then we have a tower $V_{m,0} \subset V_{m,1} \subset \cdots \subset V_{m,n}$ with $V_{m,0} = H^0(S_m, \mathcal{O}_{S_m})$. We put $V_{m,\infty} = \bigcup_n V_{m,n}$ and $T_{m,\infty} = \varprojlim_n T_{m,n}$. Similar to (4.13), we put

$$\mathcal{V} = \varinjlim_m V_{m,\infty} \text{ and } V = \varprojlim_m V_{m,\infty}. \tag{8.2}$$

On these modules, $GL_g(\mathbb{Z}_p)$ acts naturally, and for a closed subgroup $H \subset GL_g(\mathbb{Z}_p)$, we have for the submodules fixed by H:

$$\mathcal{V}^H = \varinjlim_m V_{m,\infty}^H \text{ and } V^H = \varprojlim_m V_{m,\infty}^H. \tag{8.3}$$

Indeed, $T_{\infty,\infty}/H = \varprojlim_m T_{m,m}/H_m$ for the finite image $H_m \subset GL_g(\mathbb{Z}/p^m\mathbb{Z})$ of H, and $T_{m,m}/H_m$ is the geometric quotient of $T_{m,m}$ by a finite constant group H_m (cf. [GME] Proposition 1.8.4 (b) or [SGA] V.1), since $T_{\infty,\infty} = \varprojlim_m T_{m,\infty}$ is a formal scheme which is an étale covering of S_∞ with $T_{\infty,\infty}/GL_g(\mathbb{Z}_p) = S_\infty$. This shows the second identity. The first identity is plain.

8.1.2 Rational Representation and Vector Bundles

For a given ring A or a sheaf of rings A over a scheme, we look at the flag variety $\mathcal{F}_{/A} = \mathrm{Flag}_{B/A}$ of all maximal flags in A^g defined in Section 6.1.3. We write $B \subset GL(g)$ for the upper triangular Borel subgroup. Let N be its unipotent radical, and put $T = B/N$ for the torus. Note that $T \cong B/N \cong {}^tB/{}^tN$ for the Borel subgroup ${}^tB$ opposite to B and its unipotent radical ${}^tN$. Then $\mathcal{F} \cong GL(g)/B$. We define $\mathcal{H}_{/A} = \mathrm{Flag}_{N/A} = GL(g)/N$. Write $\mathbf{1}$ for the origin of $\mathcal{H}$ represented by the coset N. Each schematic representation of T is completely reducible (see the following description of schematic representations and [RAG] I.2 and [GME] 1.6.5), and we have $H^0(\mathcal{H}, \mathcal{O}_{\mathcal{H}}) = \bigoplus_{\kappa \in X(T)} R_A[\kappa]$ for the space $R_A[\kappa]$ of weight κ. Here

$$R_A[\kappa] = \left\{ f : GL(g)/N \to \mathbf{A}^1 \in \Gamma(GL(g)/N, \mathcal{O}_{GL(g)/N}) \middle| f(ht) = \kappa(t)f(h) \right\}$$

for $t \in T$ and $\mathbf{A}^1 = Spec(A[T])$ (the one-dimensional affine space).

The pullback action of $GL(g)$ on $R_A[\kappa]$, $f(x) \mapsto \rho(h)f(x) = f(h^{-1}x)$, gives a representation $\rho = \rho_\kappa$ such that $R_A[\kappa]^N \cong A$ on which T acts by $-w_0\kappa$. Here w_0 is the longest element of the Weyl group of T represented by antidiagonal matrix $w_0 = (\delta_{i,g+1-j})_{ij}$ modulo T. The dual $R_A[\kappa]^* = \mathrm{Hom}_A(R_A[\kappa], A)$ is the universal representation of highest weight κ (cf. [RAG] II.2.13). The weight κ is *positive* (or dominant) with respect to B if $R_A[\kappa] \neq 0$, and the set of all positive weights $\kappa \in X(T)$ forms a cone $X_+(T)$ of positive weights.

To describe the notion of the *universal representation*, we recall schematic representations (cf. [GME] 1.6.5). We regard an A-projective module M as a covariant functor from A-ALG into the category of abelian groups by $\underline{M}(A') = M \otimes_A A'$. For a group scheme $G_{/A}$ and an A-projective module of finite type

M, M is called a *schematic representation* if we have a morphism of covariant functors $G \times \underline{M} \to \underline{M}$ such that the induced morphism $G(A') \times \underline{M}(A') \to \underline{M}(A')$ is an action of the group $G(A')$ on the A'-module $\underline{M}(A') = M \otimes_A A'$.

We take $G = GL(g)_{/A}$ and the schematic representation $A(\kappa)$ of $B_{/A}$ given by a character $\kappa \in X(T)$ regarded as a character of B by composition: $B \twoheadrightarrow T \xrightarrow{\kappa} \mathbb{G}_m$. Then $R_A[\kappa]^*$ satisfies the following universal property for schematic representations M of $GL(g)_{/A}$,

$$\mathrm{Hom}_{GL(g)}(R_A[\kappa]^*, M) \cong \mathrm{Hom}_B(A(\kappa), M) \quad \text{canonically.}$$

A proof and a construction of the above isomorphism is given in [RAG], II.2.13, whose outline is as follows. For a schematic $GL(g)_{/A}$-module M,

$$\begin{aligned}\mathrm{Hom}_{GL(g)}(R_A[\kappa]^*, M) &\cong \mathrm{Hom}_{GL(g)}(M^*, R_A[\kappa]) \\ &\cong \mathrm{Hom}_B(M^*, A(\kappa^{-1})) \cong \mathrm{Hom}_B(A(\kappa), M).\end{aligned}$$

The middle equality is given by the Frobenius reciprocity (e.g., [MFG] 2.1.6). The $GL(g)$-module $R_A[\kappa]^*$ is generated by a B-stable line of weight κ, and any schematic $GL(g)$-module generated by a B-stable line of weight κ is a surjective homomorphic image of $R_A[\kappa]^*$.

By this universality, taking $M = A(0)$, we know that the coinvariant space $R_A[\kappa]_N := H_0(N, R_A[\kappa])$ is a free A-module of rank 1 on which T acts by $-\kappa$, and there is a unique N-invariant linear form $\ell_{can} : R_A[\kappa] \to A$ (up to an A-unit multiple), which generates $(R_A[\kappa]^*)^N$. We can normalize ℓ_{can} so that it is the evaluation of $\phi \in R_A[\kappa]$ at the origin $\mathbf{1} \in GL(g)/N$. Then we have a tautological embedding $R_A[\kappa] \hookrightarrow \Gamma(GL(g)/N, \mathcal{O}_{GL(g)/N})$ given by

$$\phi \mapsto \{h \mapsto \ell_{can}(\rho(h^{-1})\phi)\}.$$

If $h^{-1} \in M_{g\times g/A}$ for the $g \times g$ matrix algebra $M_{g\times g/A}$ (regarded as a ring scheme over A) as a multiplicative semi-group scheme, the action of $\rho(h)$ ($h \in M_{g\times g}(A')$) is well-defined on $R_{A'}[\kappa]$ for any A-algebra A'.

In [RAG], a slightly different module is considered:

$$H^0(\kappa) = \mathrm{Ind}_{{}^tB}^{GL(g)} A(-\kappa) = \left\{ f : GL(g)/{}^tN \to \mathbf{A}^1 \middle| f(xt) = t^{-\kappa} f(x)\ \forall t \in T \right\}.$$

The action of $h \in GL(g)$ is given by $hf(x) = f(h^{-1}x)$. In this context, $-\kappa$ is a positive weight with respect to ${}^tB$, and the $H^0(\kappa)^N$ contains the highest weight vector. Using conjugation by w_0, we can remove the use of the lower triangular Borel subgroup ${}^tB$, but we need to modify the results of [RAG] accordingly when we quote them (this is done without further warning).

Suppose that A is a p-adic ring. Let $f \in R_{A/p^mA}[\kappa]$. By definition, f induces a function on $GL_g(\mathbb{Z}_p)$ by $f(h) = \rho_\kappa(h^{-1})f(1)$, and $h \mapsto \ell_{can} \circ f(h)$ is an element in $R_{A/p^mA}[\kappa]$ by tautology. This shows the following fact.

$$\textit{We have a canonical map } R_A[\kappa] \to \mathcal{C}(GL_g(\mathbb{Z}_p)/N(\mathbb{Z}_p), A)[\kappa], \qquad \text{(c)}$$

which is injective if A is flat over $\mathbb{Z}_p$. Here $\mathcal{C}(GL_g(\mathbb{Z}_p)/N(\mathbb{Z}_p), A)$ is the space of (p-adic) continuous functions with values in A on $GL_g(\mathbb{Z}_p)/N(\mathbb{Z}_p)$, and $[\kappa]$ indicates the κ-eigenspace under the right multiplication $g \mapsto gt$ by $t \in T(\mathbb{Z}_p)$ for $g \in GL_g(\mathbb{Z}_p)/N(\mathbb{Z}_p)$. The cokernel of the map (c) is large, because

1. the image is made up of polynomial functions of the coordinates of $GL(g)$;
2. $\mathcal{C}(GL_g(\mathbb{Z}_p)/N(\mathbb{Z}_p), A)$ contains all continuous functions of the quotient p-adic space $GL_g(\mathbb{Z}_p)/N(\mathbb{Z}_p)$ into A.

We let $h \in GL_g(\mathbb{Z}_p)$ act on $\mathcal{C}(GL_g(\mathbb{Z}_p)/N(\mathbb{Z}_p), A)$ by $h\phi(x) = \phi(h^{-1}x)$, and this representation of $GL_g(\mathbb{Z}_p)$ is called the *continuous induction* of the trivial representation of $N(\mathbb{Z}_p)$ to $GL_g(\mathbb{Z}_p)$ with coefficients in the p-adic ring A. When A is a finite ring, the space $\mathcal{C}(GL_g(\mathbb{Z}_p)/N(\mathbb{Z}_p), A)$ is equal to the space of locally constant functions $\mathcal{LC}(GL_g(\mathbb{Z}_p)/N(\mathbb{Z}_p), A)$, and we use the symbol $\mathcal{LC}$ instead of $\mathcal{C}$ when A is finite.

8.1.3 Weight of Automorphic Forms and Representations

Let $\underline{\omega}_m$ be the locally free sheaf $P \otimes_{\mathbb{Z}_p} \mathcal{O}_{S_m} = P_m \otimes_{\mathbb{Z}_p} \mathcal{O}_{S_m}$ on S_m of rank g. We define a coherent sheaf $\underline{\omega}_{m,n}$ on $T_{m,n}$ by $(\pi_{m,n}^* P_m) \otimes_{\mathbb{Z}} \mathcal{O}_{T_{m,n}}$ ($\underline{\omega}_m = \underline{\omega}_{m,0}$ for $T_{m,0} = S_m$). On $T_{m,n}$ with $n \geq m$, we have the universal isomorphism

$$I_{can} : \pi_{m,n}^* P_n \cong (\mathbb{Z}/p^n\mathbb{Z})^g;$$

so we have an action of $GL_g(\mathbb{Z}_p) \subset \mathrm{Gal}(T_{m,\infty}/S_m)$ on $\pi_{m,n}^* P_m$, and

$$\omega_{can} = I_{can} \otimes \mathrm{id} : \underline{\omega}_{m,n} \cong \mathcal{O}_{T_{m,n}}^g$$

is an isomorphism. Recall $\mathcal{H}_{/A} := \mathrm{Flag}_{N/A} = GL(g)/N_{/A}$. Then we can identify $\mathcal{H}_{/T_{m,n}}$ with $p_Y : Y = Y_{m,n} = GL(\underline{\omega}_{m,n})/N_{can} \to T_{m,n}$ on $V_{m,n}$ ($n \geq m$), where N_{can} is the pullback of N under ω_{can}. Thus ω_{can} induces an isomorphism $\omega_{can}^{\kappa} : p_{Y,*}(\mathcal{O}_Y[\kappa]) \cong R_{V_{m,n}}[\kappa]$. We write $\underline{\omega}_{m,n}^{\kappa}$ for the sheaf $p_{Y,*}(\mathcal{O}_Y[\kappa])$ on $T_{m,n}$. By definition, $GL_g(\mathbb{Z}/p^n\mathbb{Z})$ acts on Y on the left. The group $GL_g(\mathbb{Z}/p^n\mathbb{Z})$ acts on $\underline{\omega}_{m,n}^{\kappa}$ via the rational structure given by $\underline{\omega}_m$, and we then descend the sheaf to $\underline{\omega}_m^{\kappa}$ on S_m. In other words, for an $\mathcal{O}_{T_{m,n}}$-algebra A, $f \in H^0(Spec_{T_{m,n}}(A), \underline{\omega}_{m,n}^{\kappa})$ is a functorial rule assigning $f(X, \psi) \in R_A[\kappa]$ to $X_{/A}$ and $\psi : P_{n/X} \cong (\mathbb{Z}/p^n\mathbb{Z})^g_{/X}$. We let $h \in GL_g(\mathbb{Z}/p^n\mathbb{Z})$ act on f by $f \mapsto \{(X, \psi) \mapsto \rho(h)^{-1} f(X, h\psi)\}$. Then for any $\mathcal{O}_{T_{m,0}}$-algebra A,

$$A \mapsto H^0(GL_g(\mathbb{Z}/p^n\mathbb{Z}), H^0(Spec(A) \times_{S_m} T_{m,n}, \underline{\omega}_{m,n}^{\kappa}))$$

defines a coherent sheaf on S_m (by the Hochschild–Serre spectral sequence: [ECH] III.2.20), which we write $\underline{\omega}_m^{\kappa}$. We have

$$H^0(S_m, \underline{\omega}_m^{\kappa}) = \left\{ f \in H^0(T_{m,m}, R_{V_{m,m}}[\kappa]) \middle| f(X, h\psi) = \rho(h) f(X, \psi) \text{ for } \forall h \in GL_g(\mathbb{Z}_p) \right\}.$$

There is another description of $\underline{\omega}^{\kappa}_{m/S_m}$. Since $P_m \cong (\mathbb{Z}/p^m\mathbb{Z})^g$ on $T_{m,m}$, the action of $GL_g(\mathbb{Z}/p^n\mathbb{Z})$ on P_m extends to an action of the Galois group on $R_{\mathbb{Z}/p^m\mathbb{Z}}[\kappa]$, which determines an étale torsion sheaf P^{κ}_m over S_m. Then we have $\underline{\omega}^{\kappa}_m = \mathcal{O}_{S_m} \otimes_{\mathbb{Z}} P^{\kappa}_m$. We have $\det^k(\underline{\omega}_m) = (\bigwedge^g \underline{\omega})^{\otimes k}$ and $Sym^k(\underline{\omega}_m) = \underline{\omega}^{k\omega_1}_m$ for the first standard dominant weight ω_1.

By our definition, for each $f \in H^0(S_m, \underline{\omega}^{\kappa}_m)$,

$$\varphi(f) = \omega^{\kappa}_{can}(f) \in H^0(T_{m,m}, R_{T_{m,m}}[\kappa]),$$

which can be regarded as a functorial rule assigning each test object

$$(X_{/S_m}, \psi : P_{m/X} \cong (\mathbb{Z}/p^m\mathbb{Z})^g_{/X})$$

a value $\varphi(f)(X,\psi) \in H^0(X, R_{\mathcal{O}_X}[\kappa])$ such that $\varphi(f)(X, h\psi) = \rho(h)\varphi(f)(X,\psi)$ for all $h \in GL_g(\mathbb{Z}/p^m\mathbb{Z})$ and $\varphi(f)(Y, \phi^*\psi) = \phi^*(\varphi(f)(X,\psi))$ for a morphism $Y \xrightarrow{\phi} X$ of S_m-schemes. Similarly, $\varphi \in V_{m,n}$ is a functorial rule assigning (X,ψ) a value $\varphi(X, \psi_n : P_n \cong (\mathbb{Z}/p^n\mathbb{Z})^g) \in H^0(X, \mathcal{O}_X)$ such that $\varphi(Y, \phi^*\psi_n) = \phi^*(\varphi(X,\psi_n))$ for a morphism $Y \xrightarrow{\phi} X$ of S_m-schemes. Thus we have a natural map of $H^0(T_{m,m}, R_{T_{m,m}}[\kappa])$ into $V_{m,m}$ associated with each linear form $\ell \in R_{V_{m,m}}[\kappa]^*$. The map associates $f \in H^0(T_{m,m}, R_{T_{m,m}}[\kappa])$ with a rule: $(X,\psi) \mapsto \ell(\varphi(f)(X,\psi))$, which is a matrix coefficient of $\varphi(f)(X,\psi)$.

We let $h \in GL_g(\mathbb{Z}_p)$ act on test objects by $(X,\psi) \mapsto (X, h\psi)$. In this way, we identify $GL_g(\mathbb{Z}_p)$ with a subgroup of $\mathrm{Aut}(T_{m,\infty/S_m})$. For the Borel subgroup $B \subset GL(g)$, we put $T^B_{m,n}$ for the quotient $T_{m,n}/B(\mathbb{Z}/p^n\mathbb{Z})$. Thus $V^B_{m,n} = H^0(T^B_{m,n}, \mathcal{O}_{T^B_{m,n}})$ is made of a functorial rule $(X,\psi_n) \mapsto \varphi(X,\psi_n) \in H^0(X,\mathcal{O}_X)$ such that $\varphi(X, b\psi) = \varphi(X,\psi)$ for all ψ and $b \in B(\mathbb{Z}_p)$. We define $V^N_{m,n}$ similarly and $T^N_{m,n}$ for the unipotent subgroup $N \subset B$. Thus we have

$$\mathcal{V}^N = \mathcal{V}^{N(\mathbb{Z}_p)},\ V^N = V^{N(\mathbb{Z}_p)},\ \mathcal{V}^B = \mathcal{V}^{B(\mathbb{Z}_p)},\ \text{and } V^B = V^{B(\mathbb{Z}_p)},$$

and for these modules, we have the identity (8.3).

Let $e_1, \ldots, e_g$ be the standard base $e_\jmath = {}^t(\overbrace{0,\ldots,0,1}^{\jmath}, 0, \ldots, 0)$ of $(\mathbb{Z}/p^n\mathbb{Z})^g$, and we consider the standard filtration

$$\mathbf{1}_n : (\mathbb{Z}/p^n\mathbb{Z})^g = L_g \supset L_{g-1} \supset \cdots \supset L_0 = \{0\}$$

given by $L_\jmath = \sum_{\imath=1}^{\jmath}(\mathbb{Z}/p^n\mathbb{Z})e_\imath$. Then $\psi_n^*\mathbf{1}_n$ gives a (full) filtration $fil = fil_{\psi_n}$ of P_n, and all full filtrations $P_n = P_n^{(g)} \supset P_n^{(g-1)} \supset \cdots \supset P^{(0)} = \{0\}$ of P_n are given in this way. Since the stabilizer of $\mathbf{1}_n$ is $B(\mathbb{Z}/p^n\mathbb{Z})$, we may regard $\varphi \in V^B_{m,n}$ as a functorial rule assigning a value $\varphi(X, fil_n) \in H^0(X,\mathcal{O}_X)$ to a test object (X, fil_n). To describe $V^N_{m,n}$ in this way, we need to bring in an isomorphism of graded modules $\phi_n : gr(fil_n) \cong \oplus_{\jmath=1}^g(\mathbb{Z}/p^n\mathbb{Z})$ inducing $\phi_n^{(\jmath)} : P^{(\jmath)}/P^{(\jmath-1)} \cong (\mathbb{Z}/p^n\mathbb{Z})$. In other words, $T^N_{m,n}$ classifies triples

(X, fil_n, ϕ_n). Since we pulled back the filtration $\mathbf{1}_n$ by ψ_n, $h \in GL_g(\mathbb{Z}_p)$ acts on $(X, fil_{\psi_n}, \phi_n)$ by $fil_n \mapsto \psi^{-1}h^{-1}\psi fil_n = (h\psi_n)^*\mathbf{1}_n$ and $\phi_n \mapsto \psi^{-1}h^{-1}\psi\phi_n$.

We consider $R_{V_{m,m}}[\kappa]$ inside $V^N_{m,m}[\kappa]$, which is the homomorphic image of $H^0(S_m, \underline{\omega}^\kappa_m)$ under $f \mapsto \ell_{can} \circ \varphi(f)$. Thus we have a natural map

$$\beta : H^0(S_m, \underline{\omega}^\kappa_m) \to V^N_{m,\infty}[\kappa], \tag{8.4}$$

where $V^N_{m,m}[\kappa]$ is the κ-eigenspace of the right action of T. The above map is injective if $m = \infty$. Define

$$R'_m = \bigoplus_{\kappa \in X_+(T), \kappa \gg 0} H^0(S_m, \underline{\omega}^\kappa_m). \tag{8.5}$$

Here $\gg$ implies "sufficiently regular" as defined in Section 5.1.3.

8.1.4 Density Theorems

We assume that the formal sheaf $\underline{\omega}_\infty$ is algebraizable; so, we have a locally free sheaf $\underline{\omega}_{/S}$ of finite rank such that $\underline{\omega} \otimes_W W_m = \underline{\omega}_m$ for all m. The sheaf $\underline{\omega}$ is determined uniquely if it exists. From $\underline{\omega}$, we can create $\underline{\omega}^\kappa_{/S}$ as $\pi_*\mathcal{O}_{\mathrm{Flag}_N(\underline{\omega})}[\kappa]$ for $\pi : \mathrm{Flag}_N \to S$ (see Section 6.1.3). The global $H^0(S, \underline{\omega})$ injects into $\varprojlim_m H^0(S_m, \underline{\omega}_m)$. We define a graded module $R' = \bigoplus_\kappa H^0(S, \underline{\omega}^\kappa)$ which is embedded into $R'_\infty = \varprojlim_m R'_m$. We call an element of $H^0(S, \underline{\omega}^\kappa_{/S})$ a *false* automorphic form of weight κ, and a *true* (or classical) automorphic form is a section in $H^0(M, \underline{\omega}^\kappa_{/M})$ for a compactification $M \supset S$ of S we specify later.

We suppose now that for all $\kappa \gg 0$, the short exact sequence

$$0 \longrightarrow \underline{\omega}^\kappa_{/S} \xrightarrow{p^m} \underline{\omega}^\kappa_{/S} \longrightarrow \underline{\omega}^\kappa_{m/S_m} \longrightarrow 0$$

gives rise to an exact sequence:

$$\text{(Hp1)} \qquad 0 \longrightarrow H^0(S, \underline{\omega}^\kappa) \xrightarrow{p^m} H^0(S, \underline{\omega}^\kappa) \longrightarrow H^0(S_m, \underline{\omega}^\kappa_m) \longrightarrow 0.$$

This condition is satisfied when S_m is affine (see [ALG] III.3.5). As we show in Theorem 8.3, (Hp1) implies

$$\text{(Hp2)} \qquad V^N_{m,\infty} = V^N/p^m V^N.$$

From (Hp1), we have

$$R'/p^m R' \cong R'_m \quad \text{and} \quad H^0(S, \underline{\omega}^\kappa)/p^m H^0(S, \underline{\omega}^\kappa) \cong H^0(S_m, \underline{\omega}^\kappa_m). \tag{8.6}$$

We now define a homomorphism $\beta(m) : R'_m \to V^N_{m,m}$ in the following way. We have a canonical isomorphism $\omega_{can} = I_{can} \otimes \mathrm{id} : \underline{\omega}_{m,m} \cong \mathcal{O}^g_{T_{m,m}}$. Then

$$\beta(m)(\sum_{\kappa\gg 0} f_\kappa) = \left\{ (X_{/T_{m,m}}, \psi) \mapsto \sum_\kappa \ell_{can}(\omega_{can}^\kappa(f_\kappa(X,\psi))) \right\}$$

for $f_\kappa \in H^0(S_m, \underline{\omega}_m^\kappa)$. Here, the image of $\beta(m)$ actually falls in $V_{m,n}^N$ because $\ell_{can} \circ \rho_\kappa(u) = \ell_{can}$ for all $u \in N(\mathbb{Z}_p)$, and $\omega_{can}^\kappa(f_\kappa) \in R'_{V_{m,m}}[\kappa]$. By construction, $\beta(n) \mod p^m = \beta(m)$ for all $n > m$. Taking the projective limit, we have $\beta(\infty) : R'_\infty \to V^N = \varprojlim_m V_{m,\infty}^N$.

Since S_m is flat over $W_m = W/p^m W$ and $T_{m,n}^N = T_{m,n}/N(\mathbb{Z}/p^n\mathbb{Z})$ is étale over S_m, $V_{m,\infty}^N$ is flat over W_m. Therefore, V^N is a W-flat $GL_g(\mathbb{Z}_p)$-module.

Remark 8.1 Flatness of $H^0(M, \underline{\omega}^\kappa)$ is subtle. Let $S = \mathfrak{M}(\mathfrak{c}, \Gamma(N))\left[\frac{1}{E}\right]$ as defined in $(\Gamma(N))$ of Section 4.1.2 for a lift E of the Hasse invariant. Assume that p is prime to N and unramified in $F/\mathbb{Q}$. We take P_n to be the maximal étale quotient of $\mathbb{X}[p^n]$ for the universal AVRM $f : \mathbb{X} \to S_m$. By Corollary 4.2, we have that $\underline{\omega} \cong f_* \Omega_{\mathbb{X}/S}$. From this, the locally free sheaf $\underline{\omega}$ extends to the locally free sheaf $\underline{\omega}_{/M} = f_* \Omega_{\overline{\mathbb{X}}/M}$ on the smooth toroidal compactification M canonically because the universal abelian scheme extends to a semi-abelian scheme $\overline{\mathbb{X}}$ over the toroidal compactification as described in Sections 4.1.4 and 6.3.6. As we have seen in Remark 4.8, $H^0(M_{/W_m}, \underline{\omega}_{/M}^k)$ is not W_m-free for $1 < m < \infty$ if k is trivial on $T(\mathbb{F}_p)$, but nontrivial on $T(\mathbb{Z}/p^m\mathbb{Z})$. However, flatness of $\underline{\omega}_{/M}^k$ implies W-freeness of $H^0(M_{/W}, \underline{\omega}_{/M}^k)$.

Since B normalizes N, we can think of the action of $T = B/N$ on V^N and the κ-eigenspace $V^N[\kappa]$ of V^N. By definition, $\beta = \beta(\infty)$ induces

$$\beta = \beta_\kappa : H^0(S, \underline{\omega}^\kappa) \hookrightarrow \varprojlim_m H^0(S_m, \underline{\omega}_m^\kappa) \to V^N[\kappa].$$

Proposition 8.2 *Suppose* (Hp1) *for S. The above map β_κ is an injection.*

Proof. Since $T_{m,n}$ is faithfully flat and étale over S_m, we may make a base change $T_{m,n/S_m}^N$ to $T_{m,n/S_m}^N \times_{S_m} T_{m,m}$, and hence we may suppose that P is constant. Then $V_{m,\infty}^N$ is made of locally constant functions on $GL_g(\mathbb{Z}_p)/N(\mathbb{Z}_p)$ with values in $V_{m,0}$. By taking the limit, V^N is the space $\mathcal{C}(GL_g(\mathbb{Z}_p)/N(\mathbb{Z}_p), V_{\infty,0})$ of continuous functions on $GL_g(\mathbb{Z}_p)/N(\mathbb{Z}_p)$ with values in $V_{\infty,0} = \varprojlim_m V_{m,0}$. Then $H^0(S, \underline{\omega}^\kappa)$ is inside the limit of global sections of $\varprojlim_m R_{V_{m,0}}[\kappa]$, which injects into $\mathcal{C}(GL_g(\mathbb{Z}_p)/N(\mathbb{Z}_p), V_{\infty,0})[\kappa]$. This shows the assertion. □

We now put, for $\beta = \beta(\infty)$,

$$D' = \beta(R')\left[\frac{1}{p}\right] \bigcap V^N = \beta(\bigoplus_{\kappa\gg 0} H^0(S, \underline{\omega}^\kappa))\left[\frac{1}{p}\right] \bigcap V^N.$$

Theorem 8.3 *Suppose* (Hp1) *for S. The inclusion $\beta = \beta(\infty) : D' \to V^N$ induces an isomorphism $D'/p^m D' \cong V^N/p^m V^N$ for all m. In other words, D' is p-adically dense in V.*

This theorem shows that (Hp1) implies (Hp2).

Proof. The injectivity of $D'/p^mD' \to V^N/p^mV^N$ follows from the definition (see [K1] Appendix III), or we can argue as follows. After the faithfully flat extension to $T_{\infty,\infty} = \varprojlim_m T_{m,m}$ over $V_{\infty,0} = \varprojlim_m V_{m,0}$, P becomes constant, the base S_0 is replaced by $T_{\infty,\infty} = S_0 \times_{S_0} T_{\infty,\infty}$, and $D'' = D' \otimes_{V_{\infty,0}} V$ is the space of polynomial functions on $\mathrm{Flag}_N = (GL(g)/N)_{/\mathbb{Q}_p}$ with values in $V[\frac{1}{p}]$ whose restriction to $GL_g(\mathbb{Z}_p)/N(\mathbb{Z}_p)$ has values in V. In this situation, we have $V^N_{/T_{\infty,\infty}} = \mathcal{C}(GL_g(\mathbb{Z}_p)/N(\mathbb{Z}_p), V)$. Then it is obvious that $D''/p^mD'' = (D'/p^mD') \otimes_{V_{\infty,0}} V$ injects into

$$V^N_{/T_{\infty,\infty}} \otimes_W W_m = \mathcal{C}(GL_g(\mathbb{Z}_p)/N(\mathbb{Z}_p), V_{m,\infty}) = (V^N/p^mV^N) \otimes_{V_{\infty,0}} V.$$

Then by the faithfully flat descent, the original map is injective.

We prove that $D'/p^mD' \to V^N/p^mV^N$ is surjective. Nakayama's lemma (e.g., [BCM] II.3.2) applied to W_m and its nilpotent ideal pW_m tells us that the surjectivity for $m > 1$ follows from the surjectivity for $m = 1$. Thus we need to prove that $D'/pD' \to V^N/pV^N$ is surjective. Since $T_{m,n}/S_m$ is étale finite, replacing S_m by $T_{m,\infty}$, we may again assume that P is constant (see [K1] Appendix III pages 364–5), because we can recover the global sections of $\underline{\omega}^\kappa$ over S_m as Galois invariants of that over $T_{m,n}$. Then

$$\mathcal{O}_{T^N_{1,n}} = \mathcal{O}_{S_1} \otimes_W W[(GL(g)/N)(\mathbb{Z}/p^n\mathbb{Z})] = \mathcal{O}_{S_1}[(GL(g)/N)(\mathbb{Z}/p^n\mathbb{Z})].$$

This shows $V^N_{1,\infty} = \mathcal{LC}((GL(g)/N)(\mathbb{Z}_p), V_{1,0})$, where $\mathcal{LC}((GL(g)/N)(\mathbb{Z}_p), V_{1,0})$ is the space of locally constant functions on the p-adic space $(GL(g)/N)(\mathbb{Z}_p)$ with values in $H^0(S_1, \mathcal{O}_{S_1}) = V_{1,0}$. Writing $V^N_{\infty,0}$ as a union of W-free modules X of finite rank, we have

$$\mathcal{LC}((GL(g)/N)(\mathbb{Z}_p), V^N_{\infty,0}) = \bigcup_X \mathcal{LC}((GL(g)/N)(\mathbb{Z}_p), X).$$

Thus we need to prove that $\mathcal{LC}((GL(g)/N)(\mathbb{Z}_p), X/pX) = \mathcal{D}_X/p\mathcal{D}_X$, where $\mathcal{D}_X$ is the space of polynomial functions of homogeneous degree $\gg 0$ (with coefficients in $K = W\left[\frac{1}{p}\right]$ on the flag manifold Flag_N) which has values in X over $GL_g(\mathbb{Z}_p)$. This last fact follows from Mahler's theorem (e.g., [LFE] Section 3.2) of the density of binomial polynomials in the space of continuous functions on $\mathbb{Z}_p$ with values in $\mathbb{Z}_p$.

Let us explain the use of Mahler's theorem in more detail. Let $\mathfrak{S}$ be the subgroup of $GL_g(\mathbb{Z}_p)$ made of permutation matrices. Then the Bruhat–Tits decomposition (e.g., [RAG] I.1.9) gives us the following disjoint unions

$$GL_g(\mathbb{Z}_p) = \bigsqcup_{w\in\mathfrak{S}} N(\mathbb{Z}_p)T(\mathbb{Z}_p)w \cdot N(\mathbb{Z}_p) \quad \text{and}$$

$$GL_g(\mathbb{F}_p) = \bigsqcup_{w\in\mathfrak{S}} N(\mathbb{F}_p)T(\mathbb{F}_p)w \cdot N(\mathbb{F}_p),$$

which can be checked by a direct matrix computation. By the above decomposition of $GL_g(\mathbb{F}_p)$, we have a polynomial function on $GL(g)$ whose reduction modulo p gives rise to a characteristic function of a given double coset $N(\mathbb{Z}_p)T(\mathbb{Z}_p)w\cdot N(\mathbb{Z}_p)$. In other words, each coset except for the one containing the antidiagonal w_0 (the longest element in $\mathfrak{S}$) is Zariski closed, we can find a polynomial P vanishing outside the coset $N(\mathbb{F}_p)T(\mathbb{F}_p)w_0\cdot N(\mathbb{F}_p)$ but nonvanishing on $N(\mathbb{F}_p)T(\mathbb{F}_p)w_0\cdot N(\mathbb{F}_p)$. Then $P_{w_0}=P^{p^N-1} \mod p$ (for a suitable positive integer N) gives a characteristic function of $N(\mathbb{F}_p)T(\mathbb{F}_p)w_0\cdot N(\mathbb{F}_p)$. Then $1-P_{w_0}$ is supported by the union of $N(\mathbb{F}_p)T(\mathbb{F}_p)w\cdot N(\mathbb{F}_p)$ with w length less than that of w_0. By induction on the length of w, we can create a polynomial P_w with coefficients in the field K of fractions of W with values in W on $GL_g(\mathbb{Z}_p)/N(\mathbb{Z}_p)$ whose reduction modulo p is a characteristic function of $N(\mathbb{F}_p)T(\mathbb{F}_p)w\cdot N(\mathbb{F}_p)$. Thus for each $\phi\in\mathcal{LC}(GL_g(\mathbb{Z}_p)/N(\mathbb{Z}_p),X/pX)$, we only need to write down ϕ restricted to $N(\mathbb{Z}_p)T(\mathbb{Z}_p)w\cdot N(\mathbb{Z}_p)/N(\mathbb{Z}_p)$ as a polynomial function in $\mathcal{D}_X/p\mathcal{D}_X$. Again it is easy to see that

$$N(\mathbb{Z}_p)T(\mathbb{Z}_p)w\cdot N(\mathbb{Z}_p)/N(\mathbb{Z}_p)\cong\mathbb{Z}_p^r\times(\mathbb{Z}_p^\times)^g$$

for a suitable integer r (basically projecting down to the coordinate of $N(\mathbb{Z}_p)T(\mathbb{Z}_p)$ modulo $w\cdot N(\mathbb{Z}_p)w^{-1}\cap TN(\mathbb{Z}_p)$). Write this coordinate function as $(x,t)=(x_1,\dots,x_r,t_1,\dots,t_g)$, where t_j are the diagonal entries of T. Then on $\mathbb{Z}_p$, any locally constant function $\phi:\mathbb{Z}_p\to X/pX$ has a unique polynomial expansion $\mathbb{Z}_p\ni z\mapsto\sum_{n\geq 0}c_n\binom{z}{n}$ with $c_n\in X/pX$ (for the nth binomial polynomial $\binom{z}{n}$) by Mahler's theorem. In particular, on $\mathbb{Z}_p^\times$, the monomial $z^{n(p-1)}$ is the constant 1 modulo p. Thus we can write any locally constant function $\phi:N(\mathbb{Z}_p)T(\mathbb{Z}_p)w\cdot N(\mathbb{Z}_p)/N(\mathbb{Z}_p)\to X/pX$ as a polynomial of (x,t), and after multiplying by a high-power $\det(t)^{n(p-1)}=(\prod_i t_i)^{n(p-1)}$, the polynomial function is in the sum of the κ-eigenspaces for the sufficiently regular κs. This finishes the proof.

We now assume that there exists a proper flat scheme $M_{/W}$ such that $S\subset M$ and that $M-S$ is a proper closed subscheme of codimension ≥ 1. We further assume that $\underline{\omega}_{/S}$ extends to M. Then $\underline{\omega}_m^\kappa$ extends automatically to M_m by the theory of flag varieties given in Section 6.1.3. By the properness of M, $H^1(M,\underline{\omega}^\kappa)$ is a W-module of finite type. Taking the projective limit with respect to m of the exact sequences

$$0\longrightarrow H^0(M,\underline{\omega}^\kappa)\otimes_W W_m\longrightarrow H^0(M_m,\underline{\omega}_m^\kappa)\longrightarrow H^1(M,\underline{\omega}^\kappa)[p^m]\longrightarrow 0,$$

we get $\varprojlim_m H^0(M_m,\underline{\omega}_m^\kappa)=H^0(M,\underline{\omega}_m^\kappa)$. Let $R_m=\bigoplus_{\kappa\gg 0}H^0(M_m,\underline{\omega}_m^\kappa)$ and $R=\bigoplus_{\kappa\gg 0}H^0(M,\underline{\omega}_{/M}^\kappa)$. Then R is p-adically dense in $R_\infty=\varprojlim_m R_m$. By definition, $R\subset R'$. Note that $\det(\underline{\omega})^{p-1}$ is trivial on S_1. Let $a\in H^0(S_1,\det(\underline{\omega})^{p-1})$ be the section corresponding to $1\in\det(\underline{\omega}_1)^{p-1}\cong\mathcal{O}_{S_1}$. We assume that a extends to M_1 so that it vanishes outside S_1. Suppose that

(E) *we have a section* $E\in H^0(M,\det(\underline{\omega})^{t(p-1)})$ *such that* $E \mod \varpi=a^t$.

By further raising the power, that is, replacing E by E^{p^m}, we may assume that $E \mod p = a^t$. Then by definition,

$$H^0(S_m, \underline{\omega}_m^\kappa) = \varinjlim_n \frac{H^0(M_m, \underline{\omega}_m^\kappa) \otimes \det(\underline{\omega}_m)^{nt(p-1)})}{E^n}.$$

We now show that $\beta(R[\frac{1}{p}]) \cap V^N$ is dense in V^N. Pick $\sum_\kappa f_\kappa \in p^m V^N$ for $f_\kappa \in H^0(S, \underline{\omega}^\kappa)$. We need to approximate $f = f_\kappa$ modulo $p^{m+1}V^N$ by an element in $H^0(M, \underline{\omega}^\kappa \otimes \det(\underline{\omega})^k)$. This section $f \in H^0(S, \underline{\omega}^\kappa)$ can be written as $f \equiv g_\ell / E^\ell \mod p^{m+1}$ for $g_\ell \in H^0(M, \underline{\omega}^\kappa \otimes \det(\underline{\omega})^\ell)$. Then for $k > \ell$, we have $f \equiv g_\ell E^{k-\ell}/E^k \mod p^{m+1}$. We may assume that $k = p^m t(p-1)$. Then as a function of (X, ψ), $E^k \mod p^{m+1}$ is a constant. Thus $f \equiv g_\ell E^{k-\ell} \mod p^{m+1} R_\infty$. This shows the density of $\beta(R\left[\frac{1}{p}\right]) \bigcap V^N$ in V^N.

Corollary 8.4 *Suppose the following conditions for S in addition to* (Hp1).

(1) *We have $S \subset M$ for a proper flat scheme $M_{/W}$ such that $S_m \subset M_m = M \otimes_W W_m$ is Zariski dense for all m;*
(2) *$\underline{\omega}$ extends to a locally free vector bundle on M of rank g;*
(3) *There exist an integer $t > 0$ and a section $E \in H^0(M, \det(\underline{\omega})^{t(p-1)})$ such that $E \mod \varpi$ is the constant section 1 generating $\det(\underline{\omega}_1)^{t(p-1)} \cong \mathcal{O}_{S_1}$;*
(4) *$M - S$ is the zero locus of the section E.*

Put $D = \beta\left(\bigoplus_{\kappa \gg 0} H^0(M, \underline{\omega}^\kappa)\right)\left[\frac{1}{p}\right] \bigcap V^N$. Then D is p-adically dense in V^N.

8.1.5 *p*-Ordinary Automorphic Forms

In addition to (Hp1) (and hence (Hp2)), we now suppose having a projector e (so $e^2 = e$) acting (continuously) on V^N, which projects down $V^N[\kappa]$ onto a W-free module of finite rank (for all $\kappa \gg 0$). We put $\mathcal{V}_N = \varinjlim_n V^N/p^n V^N = \varinjlim_n V^N_{n,\infty}$. We have $\mathcal{V}_N[\kappa] = \varinjlim_m V^N_{m,m}[\kappa]$. Since V^N is W-flat, $\mathcal{V}_N$ is p-divisible, and its direct summand $e\mathcal{V}_N$ is p-divisible. A p-divisible W-module is *of finite corank* if its Pontryagin dual module is of finite rank over W.

In practice, the projector e will be constructed so that it brings $\mathcal{V}_N[\kappa]$ down onto $eH^0(M_{/W}, \underline{\omega}^\kappa \otimes \mathbb{T}_p)$ for $\kappa \gg 0$, where $\mathbb{T}_p = \mathbb{Q}_p/\mathbb{Z}_p$. This implies

$$e\mathcal{V}_N[\kappa] = eH^0(S_{/W}, \underline{\omega}^\kappa \otimes_{\mathbb{Z}_p} \mathbb{T}_p) = eH^0(M_{/W}, \underline{\omega}^\kappa \otimes_{\mathbb{Z}_p} \mathbb{T}_p) \tag{8.7}$$

if $\kappa \gg 0$. By (Hp1), $H^0(S_{/W}, \underline{\omega}^\kappa \otimes_{\mathbb{Z}_p} \mathbb{T}_p)$ is p-divisible. By assuming (8.7), $e\mathcal{V}_N[\kappa]$ is p-divisible. Since M is proper over W, $H^0(M_{/W}, \underline{\omega}^\kappa)$ is a W-module of finite type (see [EGA] III.3.2.1), and $e\mathcal{V}_N[\kappa]$ is a p-divisible module of finite corank. In any case, we simply assume that $e\mathcal{V}[\kappa]$ is p-divisible and of finite corank for $\kappa \gg 0$.

Let $\mathcal{V}_N^*$ be the Pontryagin dual module of $\mathcal{V}_N$. Since $\mathcal{V}_N$ is a discrete $T(\mathbb{Z}_p)$-module, $\mathcal{V}_N^*$ is a compact $W[[T(\mathbb{Z}_p)]]$-module. Let $T(\mathbb{Z}_p)_p = \Gamma_T$ be the p-profinite part of $T(\mathbb{Z}_p)$; in other words, we have $T(\mathbb{Z}_p) = \Gamma_T \times \Delta$ for a

finite group Δ of order prime to p. The continuous group algebra $W[[\Gamma_T]] = \varprojlim_n W[\Gamma_T/\Gamma_T^{p^n}]]$ is isomorphic to a formal power series ring $W[[x_1,\dots,x_g]]$ sending a set of minimal generators $\{\gamma_1,\dots,\gamma_g\}$ of Γ_T to $\{1+x_1,\dots,1+x_g\}$ in $W[[x_1,\dots,x_g]]$. Thus $\mathfrak{m} = (\varpi, x_1,\dots,x_g)$ is a unique maximal ideal of $W[[\Gamma_T]]$, and $W[[\Gamma_T]]$ is P-adically complete for any closed ideal $P \subset \mathfrak{m}$ under the $\mathfrak{m}$-adic topology. For an algebraic character $\kappa \in X(T)$, the character $\kappa : \Gamma_T \to W^\times$ extends by continuity to a W-algebra homomorphism $\kappa : W[[\Gamma_T]] \to W$. Then $P_\kappa = \mathrm{Ker}(\kappa : W[[\Gamma_T]] \to W)$ is a closed ideal of $W[[\Gamma_T]]$ (again by continuity).

We fix a character $\overline{\chi} : \Delta \to \mathbb{F}^\times$ for $\mathbb{F} = W/\varpi W$ and write $\widetilde{\chi} : \Delta \to W^\times$ for the Teichmüller lift of $\overline{\chi}$. We write $X_{\widetilde{\chi}} \subset X_+(T)$ for the collection of all algebraic characters $\kappa : T \to \mathbb{G}_m$ satisfying the following two conditions:

- $\kappa \equiv \widetilde{\chi} \mod \varpi$;
- κ is sufficiently regular so that the above equation (8.7) holds.

Then $X_{\widetilde{\chi}}$ is Zariski dense in $Spec(W[[\Gamma_T]])(W)$. We write $\mathcal{V}^*_{ord}$ for $e\mathcal{V}^*_N$. Let us decompose $\mathcal{V}^*_{ord} = \bigoplus_{\widetilde{\chi}\in\widehat{\Delta}} \mathcal{V}^*_{ord}[\widetilde{\chi}]$ into the direct sum of the $\widetilde{\chi}$-eigenspaces $\mathcal{V}^*_{ord}[\widetilde{\chi}]$ under the action of Δ. The summand $\mathcal{V}^*_{ord}[\widetilde{\chi}]$ is a compact module over $W[[\Gamma_T]]$. Then by (8.7), $\mathcal{V}^*_{ord}[\widetilde{\chi}] \otimes_{W[[\Gamma_T]],\kappa} W$ is W-free of finite rank $s(\widetilde{\chi})$ for $\kappa \in X_{\widetilde{\chi}}$. This implies $(\mathcal{V}^*_{ord}[\widetilde{\chi}]/P_\kappa^m \mathcal{V}^*_{ord}[\widetilde{\chi}]) \otimes_{W[[\Gamma_T]],\kappa} W \cong W^{s(\widetilde{\chi})}$ for any $m \geq 1$. We apply Nakayama's lemma (e.g., [BCM] II.3.2) to the local ring $W[[\Gamma_T]]/P_\kappa^m$ and its nilpotent ideal P_κ and find that $(\mathcal{V}^*_{ord}[\widetilde{\chi}]/P_\kappa^m \mathcal{V}^*_{ord}[\widetilde{\chi}])$ is a $W[[\Gamma_T]]$-module of finite type with minimum number $s(\widetilde{\chi})$ of generators. Taking the projective limit with respect to m, we conclude that $\mathcal{V}^*_{ord}[\widetilde{\chi}]$ is a $W[[\Gamma_T]]$-module of finite type with minimum number $s(\widetilde{\chi})$ of generators. Since $X_{\widetilde{\chi}}$ is Zariski dense in $Spec(W[[\Gamma_T]])$, we see that $\mathcal{V}^*_{ord}[\widetilde{\chi}]$ is $W[[\Gamma_T]]$-free of rank $s = s(\widetilde{\chi})$. Indeed, take s generators $v_1,\dots,v_s$ and consider the surjection $\pi : W[[\Gamma_T]]^s \to \mathcal{V}^*_{ord}[\widetilde{\chi}]$ given by $\pi(a_1,\dots,a_s) = \sum_j a_j v_j$. After taking modulo P_κ, π reduces to an isomorphism; so, $\mathrm{Ker}(\pi) \subset P_\kappa \mathcal{V}^*_{ord}[\widetilde{\chi}]$. Since $X_{\widetilde{\chi}}$ is Zariski dense, we find $\mathrm{Ker}(\pi) \subset \bigcap_{\kappa\in X_{\widetilde{\chi}}} P_\kappa \mathcal{V}^*_{ord}[\widetilde{\chi}] = \{0\}$, getting injectivity of π.

We conclude, assuming (8.7) for the middle equality,

$$\begin{aligned}\mathrm{rank}_{W[[\Gamma_T]]}\, \mathcal{V}^*_{ord}[\widetilde{\chi}] &= \mathrm{rank}_W\, \mathcal{V}^*_{ord}[\widetilde{\chi}] \otimes_{W[[\Gamma_T]],\kappa} W \\ &= \mathrm{rank}_W (eH^0(M,\underline{\omega}^\kappa) \otimes_{\mathbb{Z}_p} \mathbb{T}_p)^* = \mathrm{rank}_W\, eH^0(M,\underline{\omega}^\kappa) \quad (8.8)\end{aligned}$$

for all $\kappa \in X_{\widetilde{\chi}}$. Therefore we get

Theorem 8.5 *Suppose* (Hp1)*, the existence of the idempotent $e : V^N \to V^N$ as above and the assumptions of* Corollary 8.4*. Then $\mathcal{V}^*_{ord}$ is a well-controlled $W[[T(\mathbb{Z}_p)]]$-projective module of finite type. If we assume* (8.7)*, this means*

$$\mathcal{V}^*_{ord} \otimes_{W[[T(\mathbb{Z}_p)]],\kappa} W \cong \mathrm{Hom}_W(eH^0(M_{/W},\underline{\omega}^\kappa_{/W}), W)$$

canonically if κ is sufficiently regular. For each $\widetilde{\chi}$-component, $\mathcal{V}^*_{ord}[\widetilde{\chi}]$ is free of finite rank over $W[[\Gamma_T]]$ for the maximal p-profinite subgroup Γ_T of $T(\mathbb{Z}_p)$.

8.1.6 Construction of the Projector e_{GL}

We construct an approximation e_{GL} of the projector e. In the paper [H02], Section 2.6, it is written: "e_{GL} is constructed using solely local data of the Galois group $\mathrm{Gal}(T_{m,\infty}/S_m) = GL_n(\mathbb{Z}_p)$, while the projector e will be constructed as $e = e_G e_{GL}$ for a global projector e_G depending on the group G." This statement is misleading. We indeed need a global input to define e_{GL}. To explain this point, let us recall from Section 5.1.2 the expanding semi-group of $GL_n(\mathbb{Q}_p)$. We introduce two subgroups $I_{B,n}$ and $I_{N,n}$ of $GL_g(\mathbb{Z}_p)$ by

$$I_{B,n} = \left\{x \in GL_g(\mathbb{Z}_p) \middle| x \mod p^n \in B(\mathbb{Z}/p^n\mathbb{Z})\right\},$$
$$I_{N,n} = \left\{x \in I_{B,n} \middle| x \mod p^n \in N(\mathbb{Z}/p^n\mathbb{Z})\right\}.$$

Let $\mathrm{diag}[X_1, \ldots, X_j]$ denote the diagonal matrix with diagonal blocks X_1 to X_j from the top. We recall the subsemi-group of $GL(g)$ given by

$$D = D_{GL(g)} = \left\{\mathrm{diag}[p^{e_1}, \ldots, p^{e_g}] \middle| e_1 \le e_2 \le \cdots \le e_g\right\}.$$

Then $\Delta_n^? = I_{?,n} D I_{?,n}$ $\Delta_\infty^N = N(\mathbb{Z}_p) D N(\mathbb{Z}_p)$ and $\Delta_\infty^B = B(\mathbb{Z}_p) D B(\mathbb{Z}_p)$ are semi-groups, and we call them expanding semi-groups (see Section 5.1.2). If confusion is unlikely, we simply write Δ for one of these semi-groups.

The global input we need comes from the fact that $T_{m,n}$ in our application classifies not just trivializations of P_n but abelian varieties X with $X[p^n]^{et} \cong P_n$. In other words, each $g \in GL_n(\mathbb{Q}_p)$ acts on S_m by an appropriate isogeny of classified abelian varieties, and it acts not only on the étale quotient of the p-divisible group of the abelian variety but also on the connected component of the p-divisible group. The action changes the isomorphism class of the abelian varieties, and hence it acts on S_m through endomorphisms (not necessarily through automorphisms). Since at this point we do not assume that S_m classifies abelian varieties, *we instead assume having such an action of the expanding semi-group* (as defined below) on S_m which is at worst radiciel mod p; so, it does not affect the étale trivialization P_n. This action $\delta : S_m \to S_m$ sends an S_m-scheme $X \xrightarrow{f} S_m$ to $\delta \cdot X = X \times_{S_m,\delta} S_m$.

We consider the following triples

$$\mathcal{X} = (X_{/S_m}, fil_n, \phi_n : \oplus_{j=1}^g \mathbb{Z}/p^n\mathbb{Z} \cong gr(fil_n)),$$

where $fil_n : P_{n/X} = P_n^{(g)} \supset P_n^{(g-1)} \supset \cdots \supset P_n^{(0)} = \{0\}$ with $\mathbb{Z}/p^n\mathbb{Z} \overset{\phi_j}{\cong} P_n^{(j)}/P_n^{(j-1)}$ for $j = 1, \ldots, g$. If P is constant, the space classifying the above test objects over S_m is given by $T_{m,n}^N = T_{m,n}/N(\mathbb{Z}/p^n\mathbb{Z})$. Similarly, the classifying space of couples (X, fil_n) over S_m is given by $T_{m,n}^B = T_{m,n}/B(\mathbb{Z}/p^n\mathbb{Z})$. On test objects over $T_{m,n}$, we have a natural action of $h \in GL_n(\mathbb{Z}_p)$ given by $(X, \psi) \mapsto (X, h\psi)$. Writing $fil_n = \psi^{-1}\mathbf{1}_n$, we then see that $\psi^{-1}h^{-1}\mathbf{1}_n = \psi^{-1}\mathbf{1}_n = h \cdot fil_n$. Thus the Galois action on filtrations is given by $h \cdot fil = (\psi^{-1}h^{-1}\psi)P_n^{(j)}$ and $h \cdot \phi = (\psi^{-1}h^{-1}\psi)\phi$, where

$\psi : P_n \cong (\mathbb{Z}/p^n\mathbb{Z})^g$ such that $\psi^*(\mathbf{1}_n, \mathrm{id}) = (fil_n, \phi_n)$ for the standard identification $\mathrm{id} : gr(\mathbf{1}_n) \cong (\mathbb{Z}/p^n\mathbb{Z})^g$. Thus these test objects are always invariant under $N(\mathbb{Z}_p)$. The new test objects (X, fil_n, ϕ_n) are useful in defining an isogeny action of $\delta \in \Delta$ and in constructing the idempotent e_{GL}. We assume

(d1) δ induces an isomorphism $\delta^* fil_{n/\delta \cdot X} \cong fil_{n/X}$ compatible with the action of the semi-group on fil_n (this holds if δ mod p is radiciel), where the action of $\delta \in \Delta$ on fil_n is the multiplication by δ up to scalars (as we specify later);
(d2) $h \cdot X = X$ if $h \in GL_n(\mathbb{Z}_p)$.

Let $\mathrm{diag}[x_1, \ldots, x_\jmath]$ denote the diagonal matrix whose diagonal blocks are given by square matrices x_1 to $x_\jmath$ from the top. We put $\alpha_\jmath = \mathrm{diag}[1_{g-\jmath}, p1_\jmath]$. We consider the double coset ring $R(I_{?,n}, \Delta_n^?)$ $(n = 1, 2, \ldots, \infty$ and $? = B, N)$ made of $\mathbb{Z}$-linear combinations of double cosets $I_{?,n}\delta I_{?,n}$ for $\delta \in \Delta_n^?$.

As seen in Lemma 5.2 or its proof, once a decomposition $N(\mathbb{Z}_p)\alpha_\jmath N(\mathbb{Z}_p) = \bigsqcup_\jmath N(\mathbb{Z}_p)\delta_\jmath$ is chosen, $I_{?,n}\alpha_\jmath I_{?,n}$ for any n and ? is decomposed in the same way, $I_{?,n}\alpha_\jmath I_{?,n} = \bigsqcup_\jmath I_{?,n}\delta_\jmath$. In other words, we have for $\alpha = \prod_{\jmath=1}^{g-1} \alpha_\jmath$,

$$I_{?,n+1} \backslash I_{?,n+1}\alpha I_{?,n} \cong I_{?,n} \backslash I_{?,n}\alpha I_{?,n} \cong I_{?,n+1} \backslash I_{?,n+1}\alpha I_{?,n+1}. \tag{8.9}$$

By this fact, $R(I_{?,n}, \Delta_n^?) \cong \mathbb{Z}[t_1, \ldots, t_g]$ (a polynomial ring). See Lemma 5.2 for a proof, though the semi-group D here is a proper subsemi-group of the expanding semi-group there, strictly speaking. For $\delta \in \Delta_1^B$, the action $\rho(\delta^{-1})$ is well-defined on $R_A[\kappa]$ for any p-adic ring A, because $\rho(\delta^{-1})\phi(y) = \phi(\delta y)$ for $y \in GL(g)/N$. Decompose $I_{?,n}\delta I_{?,n} = \bigsqcup_\jmath I_{?,n}\delta_\jmath$ and regarding $f \in H^0(T_{m,n}^?, \underline{\omega}^\kappa)$ as a function of test objects $\mathcal{X}_{/T_{m,\infty}}$, we define

$$f|[I_{?,n}\delta I_{?,n}](\mathcal{X}) = \sum_\jmath \rho_\kappa(\delta_\jmath^{-1}) f(\delta_\jmath \mathcal{X}), \tag{8.10}$$

where $\delta\mathcal{X} = (\delta \cdot X, \delta(\delta^* fil_n))$. Strictly speaking, the sum above is actually heuristic, because if the action of δ on S is wildly ramified, that is, radiciel (i.e., purely inseparable in characteristic p), we need to replace the sum by the trace as already described (4.18) in Section 4.1.10 (for Hilbert modular forms). We clarify this point in Section 8.3 more carefully; so, for the moment, we content ourselves with this heuristic description.

Although we have not yet specified the action of $\delta \in \Delta_n^?$ on $\delta^* fil_n$, if it is done, then the operator is well defined independently of the choice of $\delta_\jmath$ because for $u \in I_{?,n}$ we have

$$\rho((u\delta_\jmath)^{-1}) f(u\delta_\jmath \mathcal{X}) = \rho((\delta_\jmath)^{-1})\rho(u)^{-1} f(u\delta_\jmath \mathcal{X}) = \rho(\delta_\jmath^{-1}) f(\delta_\jmath \mathcal{X}).$$

Furthermore, by (8.9),

$$\text{for } f \in H^0(T_{m,n}^B, \underline{\omega}^\kappa),\ f|t(p)^{n-1} \in H^0(T_{m,1}^B, \underline{\omega}^\kappa), \tag{8.11}$$

where $t(p) = \prod_{j=1}^{g-1} t_j$. When P_n comes from a universal abelian scheme, we have a natural isogeny action on test objects, and in this way, we can define Hecke operators on $H^0(S_m, \underline{\omega}^\kappa)$.

Since fil_n is an element of the flag variety of $(\mathbb{Z}/p^n\mathbb{Z})^g$, to study the action of Δ on filtrations, we study general flag varieties. For each commutative ring A, we consider the free module $L = L(A) = A^g$ and the flag space

$$y(A^g) = \left\{(L_i)\middle| L = L_g,\ L_i \supset L_{i-1},\ L_i/L_{i-1} \cong A \text{ for } i = 1, \cdots, g\right\}, \quad (8.12)$$

$$\mathcal{Y}(A^g) = \left\{(L_i, \phi_i)\middle|(L_i) \in y(A),\ \phi_i : A \cong L_i/L_{i-1} \text{ for } i = 1, \ldots, g\right\}. \quad (8.13)$$

We can extend the above definition to P_n over S_m. Define for T/S_m,

$$y(P_n)_{/T} = \left\{(P^{(i)}_{n/T})\middle| P_n = P^{(g)}_n,\ P^{(i)}_n \supset P^{(i-1)}_n,\ \frac{P^{(i)}_n}{P^{(i-1)}_n} \cong \frac{\mathbb{Z}}{p^n\mathbb{Z}}\ (1 \le i \le g)\right\},$$

$$\mathcal{Y}(P_n)_{/T} = \left\{(P^{(i)}_{n/T}, \phi_i)\middle|(P^{(i)}_n) \in y(P_n)_{/T},\ \phi_i : \frac{\mathbb{Z}}{p^n\mathbb{Z}} \cong \frac{P^{(i)}_n}{P^{(i-1)}_n}\ (1 \le i \le g)\right\}.$$

After a finite étale extension to T/S_m, the spaces $y(P_n)$ and $\mathcal{Y}(P_n)$ get isomorphic to $y((\mathbb{Z}/p^n\mathbb{Z})^g) \times_{S_m} T$ and $\mathcal{Y}((\mathbb{Z}/p^n\mathbb{Z})^g) \times_{S_m} T$, respectively. Writing the standard base of L as $e_1, \ldots, e_g$, we define $\mathbf{1} = (\sum_{j=1}^{i} Ae_i, \phi_i = \mathrm{id}) \in \mathcal{Y}(A^g)$, which we call the origin. We may let $GL_g(A)$ act on $\mathcal{Y}(A^g)$ and $y(A^g)$ by $x((L_i), \phi_i) = (xL_i, x \circ \phi_i)$. Then $GL_g(A)/N(A) \cong \mathcal{Y}(A^g)$ by $xN(A) \mapsto x\mathbf{1}$. Now we assume A to be a p-adic ring. We then define

$$\mathcal{Y}_n(A^g) = \{(L_i, \phi_i)|(L_i/p^nL_i) = \mathbf{1} \in y(A/p^nA)\}.$$

Similarly, we define $\mathcal{Y}_{n/T}(P_{n'})$ for $n \le n' \le \infty$. We note that $\mathcal{Y}_n(\mathbb{Z}_p^g) \cong I_{N,n}/N(\mathbb{Z}_p)$ via $x\mathbf{1} \leftrightarrow x$ and similarly $y_n(\mathbb{Z}_p^g) \cong I_{B,n}/N(\mathbb{Z}_p)$. So we have the conjugate action of Δ on these spaces introduced in Section 5.1.3.

We now write down explicitly the conjugate action of the semi-group Δ on $\mathcal{Y}_n(A^g)$. Since $y(\mathbb{Z}_p^g) = y(\mathbb{Q}_p^g)$ (because $y = \mathrm{Flag}_B$ is projective), the group $GL_g(\mathbb{Q}_p)$ acts naturally on $y(\mathbb{Z}_p^g)$. This action is described as follows. Take $x \in GL_g(\mathbb{Q}_p)$. Then $x(L_i) = (xL_i \otimes \mathbb{Q}_p \cap L(\mathbb{Z}_p)) \in y(\mathbb{Z}_p^g)$. We write $x(L_i) = (x \cdot L_i)$, that is, $x \cdot L_i = xL_i \otimes \mathbb{Q}_p \cap L(\mathbb{Z}_p)$. We claim

Lemma 8.6 *Let $d = \mathrm{diag}[p^{e_1}, \ldots, p^{e_g}] \in D_{GL(g)}$. For $(L_i, \phi_i) \in \mathcal{Y}_n(\mathbb{Z}_p^g)$, $[d]_i : L_i/L_{i-1} \to \delta \cdot L_i/\delta \cdot L_{i-1}$ given by $(\ell \mod L_{i-1}) \mapsto (p^{-e_i}d\ell \mod L_{i-1})$ for $\ell \in L_i$ is a surjective isomorphism.*

Proof. The $\mathbb{Z}_p$-module L_i is generated by L_{i-1} and a vector ℓ whose first $i-1$ entries vanish. Then $d\ell$ is divisible by p^{e_i} because its bottom $g-i$ entries are divisible by $p^{e_{i+1}}$ because of $e_g \ge \cdots \ge e_{i+1} \ge e_i$. From this, the linear map $(d)_i : (\mathbb{Q}_pL_i/\mathbb{Q}_pL_{i-1}) \to d(\mathbb{Q}_pL_i)/d(\mathbb{Q}_pL_{i-1})$ induced by d is divisible by p^{e_i} on $L_i/L_{i-1} \subset (\mathbb{Q}_pL_i/\mathbb{Q}_pL_{i-1})$. By definition, we have $d \cdot L_i = L_i$, and $p^{-e_i}(d)_i : L_i/L_{i-1} \to L_i/L_{i-1} = d \cdot L_i/d \cdot L_{i-1}$ is a surjective isomorphism. □

We now define an action of the semi-group Δ_n^B on $\mathcal{Y}_n(\mathbb{Z}_p^g)$. Since $u \in I_{B,n}$ acts naturally on flags, multiplication by u gives rise to an isomorphism $[u]_i : L_i/L_{i-1} \cong u \cdot L_i \cdot L_{i-1}$. Defining a linear map $[\delta]_i : L_i/L_{i-1} \cong \delta \cdot L_i/\delta \cdot L_{i-1}$ by $[u]_i \circ [d]_i \circ [u']_i$, the above action of $D_{GL(g)}$ on $\mathcal{Y}_n(\mathbb{Z}_p^g)$ extends to an action of the semi-group Δ_n^B. In particular, we have $\delta \cdot (L_i, \phi_i) = (\delta \cdot L_i, [\delta]_i \circ \phi_i)$ for $(L_i, \phi_i) \in \mathcal{Y}_n(\mathbb{Z}_p^g)$.

For a given $\mathcal{X} = (X, fil) = (X, \psi^{-1}\mathbf{1})$, ψ brings fil to $\mathbf{1}$, and hence the action of Δ_n^B defined on the neighborhood of $\mathbf{1}$ (after conjugation by ψ) is enough to get an association $\mathcal{X} \mapsto \{\delta_J \mathcal{X}\}$. By this, after extending base scheme S_m (e.g., to $T = T_{m,\infty}$) to trivialize $\mathcal{Y}(P)$, we have an action of Δ_n^B on $\mathcal{Y}(P)$. However, this is sufficient to define the Hecke operators $[I_{?,n}\delta I_{?,n}]$ acting on $H^0(T_{m,n/S_m}, \underline{\omega}^\kappa)$ for the following reason. After extending the scalar, define $f|[I_{?,n}\delta I_{?,n}]$ by (8.10). The formation of $f|[I_{?,n}\delta I_{?,n}]$ commutes with the base change; in other words, it commutes with the Galois action of the base $GL_g(\mathbb{Z}_p) \subset \mathrm{Gal}(T/S_m)$; so, $f|[I_{?,n}\delta I_{?,n}]$ is actually defined over the original base scheme $T^?_{m,n/S_m}$. This justifies the contraction property (8.11).

Write N for $N(\mathbb{Z}_p)$ simply. Let $t_j = N\alpha_j N$ in $R(N, \Delta_\infty^N)$ with $\alpha_j = \mathrm{diag}[1_j, p1_{g-j}]$, and define $t(p) = \prod_{j=1}^g t_j$. Let $\alpha = \prod_{j=1}^{g-1} \alpha_j$. By computation, we see easily that α^n contracts $y_n(\mathbb{Z}/p^{n+1}\mathbb{Z})$ to the origin $\mathbf{1}_{n+1}$.

Identifying $y_1(\mathbb{Z}/p^n\mathbb{Z})$ with $I_{B,1}/I_{N,n}$, if the filtration fil_n corresponds to $x \in I_{B,n}$, then the filtration is given by $\sum_{j=1}^{\iota}(\mathbb{Z}/p^n\mathbb{Z})x_j$ for the jth column vector x_j of x. Choose a representative set $N(\mathbb{Z}_p)\alpha N(\mathbb{Z}_p) = \bigsqcup_{u\in R} N(\mathbb{Z}_p)\alpha_u$. Then we have $I_{B,n+1}\alpha I_{N,n} = \bigsqcup_{u\in R} I_{B,n+1}\alpha_u$, and $\alpha_u x^{-1} = x_u^{-1}\alpha_{u'}$ for some $u' \in R$. This $x^{-1} \mapsto x_u^{-1}$ coincides with the action of α_u on the flag variety $\mathcal{Y}_1((\mathbb{Z}/p^n\mathbb{Z})^g)$ if one identifies elements in $I_{B,1}$ with a flag. Here we need to use x^{-1} instead of x, because the action of $h \in GL_g(\mathbb{Z}_p)$ on filtrations is given by $fil \mapsto h \cdot fil = (\psi^{-1}h^{-1}\psi)fil$ as already remarked. The element x_u gives rise to a couple $(\alpha_u \cdot X, fil_{n+1,u} = \alpha_u(\alpha_u^* fil_{n+1}))$, which is uniquely determined independently of the choice of α_u. We then define for $f \in V_{m,n+1}^B$ $f|\mathbf{t}(p)(X, fil_n) = \sum_{u\in R} f(\alpha_u \cdot X, fil_{n+1,u})$. Similarly, if we start from $f \in V_{m,n}^N$, by the same process, we get $x_u \in I_{N,n}/N(\mathbb{Z}_p)$ corresponding to $(X, fil_{u,n}, \phi_u)$, because we still have $I_{N,n}\alpha I_{N,n} = \bigsqcup_{u\in R} I_{N,n}\alpha_u$. Define

$$f|\mathbf{t}(p)(X, fil_n, \phi_n) = \sum_{u\in R} f(\alpha_u \cdot X, fil_{u,n}, \phi_u) \tag{8.14}$$

and define the idempotent e_{GL} by $e_{GL} = \lim_{n\to\infty} \mathbf{t}(p)^{n!}$ whenever it is well-defined. As we have seen (5.10), $\kappa(\alpha)\mathbf{t}(p) = t(p)$ on $H^0(S, \underline{\omega}^\kappa)$, because on $\underline{\omega}^\kappa$ we used the action of Δ coming from schematic induction.

8.1.7 Axiomatic Control Result

In this subsection, we describe a simple prerequisite to have the control theorem relating false automorphic forms (sections over S) to true automorphic

forms (sections over M). Later we will verify the requirement for automorphic vector bundles on Shimura varieties of PEL type.

Since in this general situation S_m is not supposed to classify anything, we cannot define Hecke operators acting on $H^0(S_m, \underline{\omega}^\kappa)$ in this generality. Anyway, we suppose a Hecke operator $\tau(p)$ acting on $H^0(S_m, \underline{\omega}^\kappa)$ and $H^0(M_m, \underline{\omega}^\kappa)$ such that $\tau(p) \equiv \mathbf{t}(p)$ on $H^0(S_1, \underline{\omega}^\kappa)$ if $\kappa \gg 0$. We define $e_0 = \lim_{n\to\infty} \tau(p)^{n!}$. Further suppose that there exist two projectors e_G acting on $V^N_{m,\infty}$ and e°_G on $H^0(M_m, \underline{\omega}^\kappa)$ for $\kappa \gg 0$ (depending on the reductive group G) such that $e_G e_{GL} = e_G$, $e^\circ_G e_0 = e^\circ_G$, and $e_G \equiv e^\circ_G \mod p$; that is, they are equal to each other on $e_G H^0(S_1, \underline{\omega}^\kappa)$. In addition to the above conditions, writing K for the field of fractions of W, we suppose the following two conditions,

(C) $e_G(E \cdot f) = E \cdot (e_G f)$ *for* $f \in H^0(S_1, \underline{\omega}^\kappa)$ *for* E *in* (E) *of* Section 8.1.3,
(F) $\dim_K e^\circ_G H^0(M_{/K}, \underline{\omega}^\kappa \otimes \det(\underline{\omega})^k_{/K})$ *is bounded independently of* k.

Let $\overline{f}_1, \overline{f}_2, \dots$ be a sequence of linearly independent elements in $e_G H^0(S_1, \underline{\omega}^\kappa)$ over W_1. Since $H^0(S_{/W}, \underline{\omega}^\kappa_{/W}) \otimes_W W_1 = H^0(S_1, \underline{\omega}^\kappa)$ (Hp1), we can lift $\overline{f}_i$ to $f_i \in H^0(S_{/W}, \underline{\omega}^\kappa_{/W})$ so that $\overline{f}_i = (f_i \mod p)$. Then for any given integer $L > 0$, we can find a sufficiently large integer m such that $E^m f_i \in H^0(M_{/W}, \underline{\omega}^\kappa_{/W})$. Since multiplication by E is an isomorphism on S_1 (by definition of $S_1 = M_1\left[\frac{1}{E}\right]$), by (C) and $e_G \equiv e^\circ_G \mod p$, the mod p sections $\{(e^\circ_G(E^m f_i) \mod p)\}_{i=1,\dots,L}$ are linearly independent over W_1 by Nakayama's lemma; so, $\{e^\circ_G(E^m f_i)\}_{i=1,\dots,L}$ are linearly independent over W. This implies

$$\dim_K e^\circ_G H^0(M_{/K}, \underline{\omega}^\kappa \otimes \det(\underline{\omega})^{t(p-1)m}_{/K}) \geq L.$$

If $\operatorname{rank}_{W_1} e_G H^0(S_1, \underline{\omega}^\kappa) = \infty$, we can take L to be arbitrarily large, which contradicts the boundedness of the dimension (F). Thus $\operatorname{rank}_{W_1} e_G H^0(S_1, \underline{\omega}^\kappa)$ has to be finite, and $\operatorname{rank}_{W_1} e_G H^0(S_1, \underline{\omega}^\kappa \otimes \det(\underline{\omega})^{k(p-1)})$ is independent of k. Thus the existence of the desired projector follows from (F), (C) and (8.7).

The condition (F) can be proven in our application, as in the proof of Theorem 4.9, via group cohomology using the (generalized) Eichler–Shimura isomorphism combined with the p-adic density of D_{cusp} in D'_{cusp} (see [H95] for such boundedness for inner and outer forms of $GL(n)$, [TiU] for inner forms of $GSp(2n)$, and Section 4.3 in the text for $GL(2)$).

The condition (C) can be proven either by q-expansion or the fact that the Hasse invariant does not change after division of an abelian variety by an étale subgroup.

Hereafter we write $H^q_{ord}(X, ?)$ for the image $e^\circ_G(H^q(X, ?))$ if $X = M$ and for the image $e_G(H^q(X, ?))$ if $X = S_m$ and $T_{m,n}$.

8.2 Deformation Theory of Serre and Tate

Recall W which is a complete discrete valuation ring of mixed characteristic with residue field $\mathbb{F}$ of characteristic p. We suppose that $\mathbb{F}$ is algebraic over

$\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. After Section 8.2.3, we suppose that $\mathbb{F}$ is an algebraic closure of $\mathbb{F}_p$. In this section, we describe deformation theory of abelian schemes over local W_m-algebras for $W_m = W/p^m W$. This theory is useful for proving the p-integrality of the Hecke operator $T(\xi)$ after the division by $\mu(\xi)$ as in (5.40). We principally follow Katz's exposition [K2].

8.2.1 A Theorem of Drinfeld

In this and the following subsection, we take a local W_m-algebra $\mathcal{B}$ as a base ring. We consider the category $\mathcal{B}$-$LR = LR_{/\mathcal{B}}$ made up of p-adically complete local $\mathcal{B}$-algebras $R = \varprojlim_n R/p^n R$ and morphisms are supposed to be p-adically continuous. For simplicity, we always assume that rings we consider are Noetherian.

Let $G : \mathcal{B}\text{-}LR \to AB$ be a covariant functor into the category AB of abelian groups. Thus if we regard G as a functor from the category of affine $\mathcal{B}$-schemes (or formal $\mathcal{B}$-schemes), it is contravariant. Suppose that, for any faithfully flat extension of finite type $R \hookrightarrow C$ of $\mathcal{B}$-algebras,

1. The group $G(R)$ injects into $G(C)$, that is, $G(R) \hookrightarrow G(C)$;
2. Let $C' = C \otimes_R C$ and $C'' = C \otimes_R C \otimes_R C$. Write $\iota_i : C \hookrightarrow C'$ $(i = 1, 2)$ two inclusions (i.e., $\iota_1(r) = r \otimes 1$) and $\iota_{ij} : C' \hookrightarrow C''$ three inclusions (i.e., $\iota_{12}(r \otimes s) = r \otimes s \otimes 1$). If $x \in G(C)$ satisfies $G(\iota_1)(x) = G(\iota_2)(x)$ and $G(\iota_{12})(y) = G(\iota_{23})(y) = G(\iota_{13})(y)$ for $y := G(\iota_1)(x)$, then x is in the image of $G(R)$.

Such a G is called an abelian sheaf on $\mathcal{B}$-LR under the $fppf$-topology (or simply an *abelian $fppf$-sheaf*). If $A_{/\mathcal{B}}$ is an abelian scheme, then

$$G(R) = A(R) = \mathrm{Hom}_S(Spec(R), A)$$

($S = Spec(\mathcal{B})$ or $Spf(\mathcal{B})$) is an $fppf$-sheaf. We call an $fppf$-sheaf $G_{/\mathcal{B}}$ a *p-divisible* group if for any $x \in G(R)$, there exists a finite faithfully flat extension C of R and a point $y \in G(C)$ such that $x = py$. If G is p-divisible of p-torsion and $G[p]_{/\mathcal{B}} = \mathrm{Ker}(p : G \to G)$ is a group scheme locally $\mathcal{B}$-free of finite rank, it is called a *Barsotti–Tate* group (cf. [CBT] I). If G comes from an abelian scheme A, it is p-divisible. This also shows that $A[p^\infty] = \bigcup_n A[p^n]$ for $A[p^n] = \mathrm{Ker}(p^n : A \to A)$ is a Barsotti–Tate group (e.g., [GME] 4.1.18).

Let $\mathcal{B}$ be a local W_m-algebra and I be an ideal of $\mathcal{B}$ such that $I^{\nu+1} = 0$ and $NI = 0$ for a power N of p. We define a new functor G_I and $\widehat{G}$ by

$$G_I(R) = \mathrm{Ker}(G(R) \to G(R/I)) \quad \text{and} \quad \widehat{G}(R) = \mathrm{Ker}(G(R) \to G(R/\mathfrak{m}_R)),$$

where $\mathfrak{m}_R$ is the maximal ideal of R. When $\widehat{G}(R) = \mathrm{Hom}_{\mathcal{B}\text{-}LR}(\mathcal{R}, R)(= G(R))$ for $\mathcal{R} = \mathcal{B}[[T_1, \ldots, T_n]]$ (i.e., $\widehat{G}_{/\mathcal{B}} = Spf(\mathcal{R})_{/\mathcal{B}}$) and the identity element $\mathbf{0}$ corresponding to the ideal $(T_1, \ldots, T_n)$, we call $\widehat{G}$ a *formal* group. If $\widehat{G}$ is formal, $G_I(R) = \{(t_1, \ldots, t_n) \in I^n\}$ by $\mathrm{Hom}_{\mathcal{B}\text{-}LR}(\mathcal{R}, R) \ni \phi \mapsto (\phi(T_1), \ldots, \phi(T_n))$.

Suppose that $\widehat{G}_{/\mathcal{B}}$ is formal. Then multiplication by $[N]$ induces a continuous algebra homomorphism $[N] : \mathcal{R} \to \mathcal{R}$. Then on the tangent space at the origin $t_G = (T_1, \ldots, T_n)/(T_1, \ldots, T_n)^2$, the multiplication by p on G (which is written as $[p] : G \to G$) induces multiplication by p, because the addition induced by the group law of G coincides with the addition of the tangent vectors (cf. [ABV] Section 11) by the uniqueness of the $\mathcal{B}$-linear structure on t_G. Thus $[N](T_i) \equiv NT_i \mod (T_1, \ldots, T_n)^2$, and $[N](G_I(R)) = G_{I^2}(R)$ because $NI = 0$. Similarly, we have $[N](G_{I^a}(R)) = G_{I^{a+1}}(R)$ inductively. This implies that $[N^\nu]G_I = G_0 = \{\mathbf{0}\}$. We get

$$G_I \subset \mathrm{Ker}([N^\nu] : G \to G) \quad \text{if } G \text{ is formal.} \tag{8.15}$$

Theorem 8.7 (Drinfeld) *Let G and H be abelian fppf-sheaves over $\mathcal{B}$-LR and I be as above, and suppose that $I^{\nu+1} = 0$. Let G_0 and H_0 be the restriction of G and H to $\mathcal{B}/I$-LR. Suppose*

(i) *G is p-divisible;*
(ii) *$\widehat{H}$ is formal;*
(iii) *$H(R) \to H(R/J)$ is surjective for any nilpotent ideal J (i.e., H is formally smooth).*

Then we have the following four assertions:

(1) *$\mathrm{Hom}_{\mathcal{B}\text{-}Gp}(G, H)$ and $\mathrm{Hom}_{\mathcal{B}/I\text{-}Gp}(G_0, H_0)$ is p-torsion-free, where the symbol "$\mathrm{Hom}_{X\text{-}Gp}$" stands for the homomorphisms of abelian fppf-sheaves over X-LR;*
(2) *The natural map, the so-called "reduction $\mod I$" from $\mathrm{Hom}_{\mathcal{B}\text{-}Gp}(G, H)$ into $\mathrm{Hom}_{\mathcal{B}/I\text{-}Gp}(G_0, H_0)$ is injective;*
(3) *For $f_0 \in \mathrm{Hom}_{\mathcal{B}/I\text{-}Gp}(G_0, H_0)$, we have a unique $\Phi \in \mathrm{Hom}_{\mathcal{B}\text{-}Gp}(G, H)$ such that $\Phi \mod I = N^\nu f_0$. We write $\widetilde{N}^\nu f$ for Φ in $\mathrm{Hom}_{\mathcal{B}\text{-}Gp}(G, H) \otimes_{\mathbb{Z}} \mathbb{Q}$;*
(4) *In order that $f = \frac{\widetilde{N}^\nu f}{N^\nu} \in \mathrm{Hom}_{\mathcal{B}\text{-}Gp}(G, H)$, it is necessary and sufficient that $\widetilde{N}^\nu f$ kills $G[N^\nu]$.*

Proof. The first assertion follows from p-divisibility, because if $pf(x) = 0$ for all x, taking y with $py = x$, we find $f(x) = pf(y) = 0$ and hence $f = 0$.

We have an exact sequence $0 \to H_I \to H \to H_0 \to 0$; so, we have another exact sequence

$$0 \to \mathrm{Hom}(G, H_I) \to \mathrm{Hom}(G, H) \xrightarrow{\mod I} \mathrm{Hom}(G, H_0) = \mathrm{Hom}(G_0, H_0),$$

which tells us the injectivity since H_I is killed by N^ν and $\mathrm{Hom}(G, H)$ is p-torsion-free.

To show (3), take $f_0 \in \mathrm{Hom}(G_0, H_0)$. By the surjectivity of $H(R) \to H_0(R/I)$, we can lift $f_0(x \mod I)$ to $y \in H(R)$. The class $y \mod \mathrm{Ker}(H \to H_0)$ is uniquely determined. Since $\mathrm{Ker}(H \to H_0)$ is killed by N^ν, $N^\nu y$ for any $x \in G(R)$ is uniquely determined; so, $x \mapsto y$ induces the functorial map $\Phi = \widetilde{N}^\nu f : G(R) \to H(R)$. This shows (3).

The assertion (4) is then clear from p-divisibility of G. The uniqueness of f follows from the p-torsion-freeness of $\mathrm{Hom}(G,H)$. □

8.2.2 A Theorem of Serre–Tate

Let $\mathcal{A}_{/\mathcal{B}}$ be the category of abelian schemes defined over $\mathcal{B}$. We consider a category $Def(\mathcal{B},\mathcal{B}/I)$ of triples $(A_0,\mathcal{G},\epsilon)$, where A_0 is an abelian scheme over $\mathcal{B}_0=\mathcal{B}/I$, a Barsotti–Tate group $\mathcal{G}$ over $\mathcal{B}$ and $\epsilon:\mathcal{G}_0\cong A_0[p^\infty]$. We have a natural functor $\mathcal{A}_{/\mathcal{B}}\to Def(\mathcal{B},\mathcal{B}_0)$ given by

$$A\mapsto (A_0=A \mod I, A[p^\infty], \mathrm{id}).$$

Theorem 8.8 (Serre–Tate) *The above functor $\mathcal{A}_{/\mathcal{B}}\to Def(\mathcal{B},\mathcal{B}_0)$ is an equivalence of categories.*

Proof. By Drinfeld's theorem applied to $A[p^\infty]$ and A (both abelian $fppf$-sheaf), the functor is fully faithful.

It is known that we can lift A_0 to an abelian scheme Y over $\mathcal{B}$. This follows from the deformation theory of Grothendieck (e.g., [CBT] 2.8.1). Here is a heuristic argument to find $Y_{/\mathcal{B}}$ under the assumption that $\mathcal{B}_0$ is a finite field. In this case, by a theorem of Tate (see [T1], [D], and [ABV] Section 22), A_0 has complex multiplication by a semi-simple commutative algebra M (with $2\cdot\dim A_0=\dim_{\mathbb{Q}} M$) whose simple components are all CM fields with its complex conjugation induced by a Rosati involution. By the theory of abelian varieties with complex multiplication, A_0 can be lifted to an abelian scheme Y over $\mathcal{B}$ with complex multiplication (this is a special case of Theorem 7.1 taking $D=V=M$). Thus we have an isomorphism $\alpha_0^{(p)}:Y_0[p^\infty]\to A_0[p^\infty]$, and we have a unique lifting (by Drinfeld's theorem) $f=\widetilde{N^\nu\alpha^{(p)}}:Y[p^\infty]\to\mathcal{G}$ of $N^\nu\alpha_0^{(p)}$. Since the special fiber is an isogeny having inverse $(\alpha_0^{(p)})^{-1}$, f is an isogeny, whose (quasi) inverse is the lift of $N^\nu(\alpha_0^{(p)})^{-1}$. Thus $\mathrm{Ker}(f)$ is a finite flat group subscheme of Y. The geometric quotient of Y by a finite flat group subscheme exists (see [ABV] Section 12) and is an abelian scheme over $\mathcal{B}$. Then dividing Y by $\mathrm{Ker}(f)$, we get the desired $A_{/\mathcal{B}}\in\mathcal{A}_{/\mathcal{B}}$. □

Suppose $\mathcal{B}_0$ is a field of characteristic p. As described in [CBT], IV.2.5 and V.1.4, we can attach a crystal $(\mathbb{D}_{\mathcal{B}}(\mathcal{G}_0)\supset F^1)$ to the data $(A_0,\mathcal{G},\varepsilon)$ formed by a $\mathcal{B}$-free module $\mathbb{D}_{\mathcal{B}}(\mathcal{G}_0)$ and a submodule $F^1\subset\mathbb{D}_{\mathcal{B}}(\mathcal{G}_0)$ with $\mathcal{B}$-free quotient $\mathbb{D}_{\mathcal{B}}(\mathcal{G}_0)/F^1$ such that $\mathbb{D}_{\mathcal{B}}(\mathcal{G}_0)\otimes_{\mathcal{B}}\mathcal{B}_0=\mathbb{D}_{\mathcal{B}_0}(\mathcal{G}_0)$ is the Dieudonné module of $\mathcal{G}_0$ and $F^1\otimes_{\mathcal{B}}\mathcal{B}_0$ gives $Lie(\mathcal{G}_0)\subset\mathbb{D}_{\mathcal{B}_0}(\mathcal{G}_0)$. Then we have

$$\mathrm{End}(A_{/\mathcal{B}})=\{\alpha\in\mathrm{End}_{\mathcal{B}}(A_{0/\mathcal{B}_0})|\alpha_*(F^1)\subset F^1\},$$

regarding $\mathrm{End}(A_{/\mathcal{B}})\subset\mathrm{End}(A_{0/\mathcal{B}_0})$ by Drinfeld's theorem. By this fact combined with a version of the above theorem ([CBT] V.1.10), one can verify (using (det) and D-linearity of the polarization λ_0) that $(A_0,\lambda_0,i_0,\eta^{(p)})\in\mathcal{E}^{(p)}(\mathcal{B}_0)$ for the functor $\mathcal{E}^{(p)}$ in Theorem 7.1 can be lifted to $(A,\lambda,i,\eta^{(p)})\in\mathcal{E}^{(p)}(\mathcal{B})$ (as shown in [LaR] 6.2 and [Zi]).

8.2.3 Deformation of an Ordinary Abelian Variety

Hereafter we suppose that the residue field $\mathbb{F}$ of W is an algebraic closure of $\mathbb{F}_p$. We assume the base ring $\mathcal{B}$ to be a pro-Artinian local ring with residue field $\mathbb{F}$. Write $CL_{/\mathcal{B}}$ for the category of complete local $\mathcal{B}$-algebras with residue field $\mathbb{F}$ (under the $\mathfrak{m}_{\mathcal{B}}$-adic topology). We fix an *ordinary* abelian variety $A_{0/\mathbb{F}}$. Recall that ${}^tA_{/\mathcal{B}}$ denotes the dual abelian scheme of an abelian scheme $A_{/\mathcal{B}}$ and that $\widehat{A}_{/\mathcal{B}}$ denotes the formal group associated with A. We write $\mathcal{T}A[p^\infty]^{et}$ for the Tate module of the maximal étale quotient of $A[p^\infty]$. We consider the deformation functor $\widehat{\mathcal{P}} = \widehat{\mathcal{P}}_{A_0} : CL_{/\mathcal{B}} \to SETS$ given by

$$\widehat{\mathcal{P}}_{A_0}(R) = \left[(A_{/R}, \iota_A) \middle|\ A_{/R} \text{ is an abelian scheme and } \iota_A : A \otimes_R \mathbb{F} \cong A_0\right].$$

Here $f : (A, \iota_A)_{/R} \cong (A', \iota_{A'})_{/R}$ if $f : A \to A'$ is an isomorphism of abelian schemes with the following commutative diagram,

$$\begin{array}{ccc} A \otimes_R \mathbb{F} & \xrightarrow{f_0} & A' \otimes_R \mathbb{F} \\ {\scriptstyle \iota_A}\downarrow\wr & & {\scriptstyle \iota_{A'}}\downarrow\wr \\ A_0 & = & A_0\,. \end{array}$$

Theorem 8.9 (Serre–Tate) *Let the notation and assumption be as above. Then we have*

(1) *A canonical isomorphism* $\widehat{\mathcal{P}}(R) \cong \mathrm{Hom}(\mathcal{T}A_0[p^\infty]^{et} \times \mathcal{T}^tA_0[p^\infty]^{et}, \widehat{\mathbb{G}}_m(R))$ *taking* $(A_{/R}, \iota_A)$ *to* $q_{A/R}(\cdot,\cdot)$.

(2) *The functor* $\widehat{\mathcal{P}}$ *is represented by the formal scheme*

$$\mathrm{Hom}_{\mathbb{Z}_p}(\mathcal{T}A_0[p^\infty]^{et} \times \mathcal{T}^tA_0[p^\infty]^{et}, \widehat{\mathbb{G}}_m) \cong \widehat{\mathbb{G}}_m^{g^2}.$$

(3) $q_{A/R}(x,y) = q_{{}^tA/R}(y,x)$ *under the canonical identification:* ${}^t({}^tA) = A$.

(4) *Let* $f_0 : A_{0/\mathbb{F}} \to Y_{0/\mathbb{F}}$ *be a homomorphism of two ordinary abelian varieties with the dual map* ${}^tf_0 : {}^tY_0 \to {}^tA_0$. *Then* f_0 *is induced by a homomorphism* $f : A_{/R} \to Y_{/R}$ *for* $A \in \widehat{\mathcal{P}}_{A_0}(R)$ *and* $Y \in \widehat{\mathcal{P}}_{Y_0}(R)$ *if and only if* $q_{A/R}(x, {}^tf_0(y)) = q_{Y/R}(f_0(x), y)$.

Proof. We give a sketch of the construction of $q_{A/R}$. First we prepare some facts. Let $f : A_{/R} \to Y_{/R}$ be an isogeny; so, $\mathrm{Ker}(f)$ is a finite flat group scheme over S. Pick $x \in \mathrm{Ker}(f)$, and let $\mathcal{L} \in \mathrm{Ker}({}^tf) \subset \widehat{Y}$ be a line bundle on Y with $\mathbf{0}^*\mathcal{L} = \mathcal{O}_S$ ($S = Spec(R)$). Recall that R is an Artinian $\mathcal{B}$-algebra. Thus $f^*\mathcal{L} = \mathcal{O}_A$. Cover Y by affine subsets $U_i = Spec(R_i)$ so that $\mathcal{L}|_{U_i} = \phi_i^{-1}\mathcal{O}_{U_i}$. Since $\mathbf{0}_Y^*\mathcal{L} = \mathcal{O}_S$, we may assume that $(\phi_i/\phi_j) \circ \mathbf{0}_Y = 1$. Since $f : A \to Y$ is finite, it is affine. Write $V_i = f^{-1}(U_i) = Spec(R_i)$. Then $f^*\mathcal{L}|_{V_i} = \varphi_i^{-1}\mathcal{O}_{V_i}$ with $\varphi_i = \phi_i \circ f$, and we have, regarding $x : S \to \mathrm{Ker}(f)$,

$$\frac{\varphi_i \circ x}{\varphi_j \circ x} = \frac{\phi_i \circ f \circ x}{\phi_j \circ f \circ x} = \frac{\phi_i \circ \mathbf{0}_Y}{\phi_j \circ \mathbf{0}_Y} = 1.$$

The morphisms $\varphi_i \circ x$ glue into a morphism $e_f(x, \mathcal{L}) : S \to \mathbb{G}_m$, and we get a pairing $e_f : \mathrm{Ker}(f) \times \mathrm{Ker}({}^t f) \to \mathbb{G}_m$. Since A is a $\mathrm{Ker}(f)$-torsor over Y, we have $A \times_Y A \cong \mathrm{Ker}(f) \times_S Y$. For a homomorphism $\zeta : \mathrm{Ker}(f) \to \mathbb{G}_m$, we can find a morphism $\phi : \mathrm{Ker}(f) \times_S Y \to \mathbf{P}^1$ of schemes such that $\phi(y+t) = \zeta(t)\phi(y)$ for $t \in \mathrm{Ker}(f)$. This function ϕ gives rise to a divisor D on $Y_A = Y \times_S A$. By definition $f_A^* \mathcal{L}(D) = \mathcal{O}_{A_A}$, and $e_f(x, \mathcal{L}(D)) = \zeta(x)$. Thus, over A, $e_{f/A} : \mathrm{Ker}(f)_{/A} \times \mathrm{Ker}({}^t f)_{/A} \to \mathbb{G}_m$ is a perfect pairing. Since $A \to S$ is faithfully flat, we find that the original e_f is perfect.

We apply the above argument to $f = [p^n] : A \to A$, write the pairing as e_n, and verify the following points.

(P1) $e_n(\alpha(x), y) = e_n(x, {}^t\alpha(y))$ for $\alpha \in \mathrm{End}(A_{/R})$;

(P2) Write $A_0[p^n]^\circ = \mu_{p^n}^g \subset A_0[p^n]$. Then e_n induces an isomorphism of group schemes $A_0[p^n]^\circ \cong \mathrm{Hom}({}^t A_0[p^n]^{et}, \mu_{p^n})$;

(P3) Taking the limit of the above isomorphisms with respect to n, as formal groups, we find $\widehat{A} \cong \mathrm{Hom}(\mathcal{T}^t A[p^\infty]^{et}, \widehat{\mathbb{G}}_m) \cong \mathrm{Hom}(\mathcal{T}^t A_0[p^\infty]^{et}, \widehat{\mathbb{G}}_m)$. The induced pairing we denote by $E_A : \widehat{A} \times \mathcal{T}^t A_0[p^\infty]^{et} \to \widehat{\mathbb{G}}_m$. In particular $\widehat{A} = \widehat{\mathbb{G}}_m^g$.

We now describe q. Since $R \in CL_{/\mathcal{B}}$ is a projective limit of local $\mathcal{B}$-algebras with nilpotent maximal ideal, we may assume that R is a local $\mathcal{B}$-algebra with nilpotent maximal ideal $\mathfrak{m}_R$. Then $\widehat{A}(R)$ is killed by p^{n_0} for a sufficiently large n_0 (applying Drinfeld's theorem to $I = \mathfrak{m}_R$). Thus if $n \geq n_0$, the structure of the p-divisible group $A[p^\infty]$ is uniquely determined by the extension class of

$$0 \to \widehat{A}(R) \to A[p^n](R) \xrightarrow{\pi} A_0[p^n]^{et}(R) \to 0 \tag{8.16}$$

for $n \geq n_0$, which can be rewritten as the following exact sequence:

$$0 \to \mathrm{Hom}(\mathcal{T}^t A_0[p^\infty]^{et}, \widehat{\mathbb{G}}_m)[p^n] \to A[p^n] \xrightarrow{\pi} A_0[p^n]^{et} \to 0. \tag{8.17}$$

Take $x = \varprojlim_n x_n \in \mathcal{T} A_0[p^\infty]^{et}$ for $x_n \in A[p^n]^{et}$. Lift x_n to $v_n \in A[p^n](R)$ so that $\pi(v_n) = x_n$. Then, for $\widetilde{p^n} : A[p^n] \to \widehat{A}$ (lifting p^n times the inclusion $A_0[p^n] \hookrightarrow A_0$), $q_n(x) = \widetilde{p^n} v_n \in \widehat{A}(R)$. Take the limit of q_n to get $q(x) \in \mathrm{Hom}(\mathcal{T}^t A_0[p^\infty]^{et}, \widehat{\mathbb{G}}_m(R))$. This $q(x)$ determines the extension class of (8.16) and thus the deformation $A_{/R}$ because it is determined by $(A_0, A[p^\infty])$ by the Serre–Tate theorem in the previous subsection. Then we define $q_{A/R}(x, y) = q(x)(y)$.

It is known that for any given $q(x, y)$ as above an extension (8.16) exists by the theory of Barsotti–Tate groups studied by Messing (see [CBT] Appendix). This shows assertions (1) and (2). All other assertions follow easily from (P1-3). □

8.2.4 Symplectic Case

We now fix a polarization $\lambda_0 : A_0 \to {}^t A_0$ of degree prime to p. We consider the functor $\widehat{\mathcal{P}} = \widehat{\mathcal{P}}_{A_0, \lambda_0} : CR_{/W} \to SETS$ given by

$$\widehat{\mathcal{P}}_{A_0,\lambda_0}(R) = \left[(A_{/R}, \iota_A, \lambda) \middle| (A, \iota_A) \in \widehat{\mathcal{P}}_{A_0}(R) \text{ and } \lambda \text{ induces } \lambda_0\right].$$

Here we call $f : (A, \lambda_A, \iota_A) \to (Y, \lambda_Y, \iota_Y)$ an isomorphism if $f : (A, \iota_A) \cong (Y, \iota_Y)$ and ${}^t f \circ \lambda_Y \circ f = \lambda_A$. Note that by the Drinfeld theorem, $\mathrm{End}(A_{/R})$ is torsion-free and, hence, $\mathrm{End}(A_{/R}) \hookrightarrow \mathrm{End}^{\mathbb{Q}}(A_{/R}) = \mathrm{End}(A_{/R}) \otimes_{\mathbb{Z}} \mathbb{Q}$. We write $\alpha^* = \lambda_0^{-1} \circ {}^t\alpha \circ \lambda_0$ for $\alpha \in \mathrm{End}(A_0) \otimes_{\mathbb{Z}} \mathbb{Q}$. Since $\mathrm{End}(A_{/R}) \subset \mathrm{End}(A_0)$ again by Drinfeld's theorem, the involution keeps $\mathrm{End}^{\mathbb{Q}}(A_{/R})$ stable (because on $\mathrm{End}^{\mathbb{Q}}(A_{/R})$, it is given by $\alpha^* = \lambda^{-1} \circ {}^t\alpha \circ \lambda$). The involution $\alpha \mapsto \alpha^*$ is known to be positive (see [ABV] Section 21). The polarization λ_0 induces an isomorphism $\lambda_0 : A[p^\infty]^{et} \cong {}^t A[p^\infty]^{et}$. We identify $\mathcal{T}A_0[p^\infty]^{et}$ and $\mathcal{T}^t A_0[p^\infty]^{et}$ by λ_0. Then the involution $\alpha \mapsto {}^t\alpha$ is replaced by the Rosati involution $*$. Then it is clear from the previous theorem that

$$\widehat{\mathcal{P}}_{A_0,\lambda_0}(R) \cong \mathrm{Hom}_{\mathbb{Z}_p}(Sym^2 \mathcal{T}A_0[p^\infty]^{et}, \widehat{\mathbb{G}}_m(R)),$$

where $Sym^2 \mathcal{T}A_0[p^\infty]^{et}$ is the self tensor product $\mathcal{T}A_0[p^\infty]^{et} \otimes_{\mathbb{Z}_p} \mathcal{T}A_0[p^\infty]^{et}$ modulo the subspace generated by $x \otimes y - y \otimes x$ for all $x, y \in \mathcal{T}A_0[p^\infty]^{et}$.

8.2.5 Unitary Case

We fix a division algebra D with positive involution $*$. Assume (unr) for D in Section 7.1.1. The center of D is either a CM field F ($*$ inducing complex conjugation on F) or a totally real field on which $*$ is trivial. We fix a D-module V with $*$-Hermitian alternating form $\langle\ ,\ \rangle$ satisfying conditions (L1–2) in Chapter 7. Since D is a division algebra, $V \cong D^r$; so, we identify V with the row vector space D^r by fixing a base of V over D. Under this expression, $C = \mathrm{End}_D(V)$ is the matrix algebra $M_r({}^t D)$ for the *opposite* algebra ${}^t D$ of D (i.e., ${}^t D = D$ as sets but the product $xy \in {}^t D$ of $x, y \in {}^t D$ is defined to be the product $yx \in D$). The left action of $b \in M_r({}^t D)$ on D^r is given by $bv := v \cdot b$ with matrix right multiplication $v \cdot b$. By this expression, $\langle x, y \rangle = \mathrm{Tr}_{D/\mathbb{Q}}(x s {}^t y^*)$ for an invertible matrix $s \in M_r(D)$ with ${}^t s^* = -s$, where $\mathrm{Tr}_{D/\mathbb{Q}}$ is the reduced trace $D \to F$ composed with the field trace $\mathrm{Tr}_{F/\mathbb{Q}} : F \to \mathbb{Q}$. On C, the involution ι is given by $c^\iota = s^t c^* s^{-1}$. Recall the groups $G(R) = \{x \in C \otimes_{\mathbb{Q}} R | x x^\iota \in R^\times\}$ and $U(R) = \{x \in G(R) | x x^\iota = 1\}$.

For simplicity, we suppose either $F = F_0 = \mathbb{Q}$ or that F is an imaginary quadratic field $\mathbb{Q}[\sqrt{-\delta}]$ with discriminant $0 > -\delta \in \mathbb{Z}$. Let $\varepsilon = \mathrm{diag}[1, 0, \ldots, 0] \in O_{D,p}$ under the identification $O_{D,p} = M_d(O_p)$ in (unr) of Section 7.1.1. For the involution $*$ of $M_d(F_p)$, $x \mapsto {}^t x^* = ({}^t x)^*$ is an F_0-algebra automorphism of $M_d(F_p)$; so, it is inner up to $\sigma \in \mathrm{Gal}(F/F_0)$ (by the theorem of Skolem–Noether; [BAL] VIII.10.1). Thus by choosing the identification $O_{D,p} = M_d(O_p)$ well, we may assume that $\varepsilon^* = \varepsilon$.

We first treat the case where F is imaginary. As before, we write O for the integer ring of F. Since $D_{\mathbb{R}} = D \otimes_{\mathbb{Q}} \mathbb{R}$ can be identified with $M_q(\mathbb{C})$ with $b^* = {}^t\overline{b}$ for complex conjugation $b \mapsto \overline{b}$, $S = \sqrt{-1}s \in M_r({}^t D_{\mathbb{R}}) = M_{qr}(\mathbb{C})$ is a Hermitian matrix. Thus $U(\mathbb{R})$ is the unitary group of S. Writing the signature

of S as (m,n), we find $U(\mathbb{R}) \cong U_{m,n}$ in Section 5.2.2. By the condition (det), the representation of F on $\varepsilon(Lie(A))$ for $(A,\lambda,i,\eta) \in \mathcal{P}(R)$ $(R \in W\text{-}CL)$ is $m\,\mathrm{id} + nc$ for $\mathrm{id} : O \hookrightarrow W$ and nontrivial automorphism c of F. We fix an O_D-lattice L of V such that $\langle\ \rangle$ induces a self-duality of $L_p = L \otimes_{\mathbb{Z}} \mathbb{Z}_p$.

Lemma 8.10 *Suppose* (unr) *and* (L1-2) *in* Section 7.1.1 *and that D is a division algebra with positive involution of the second kind. If we have an ordinary triple $(A_0, \overline{\lambda}_0, \eta_0^{(p)})$ sitting over a characteristic p point of $Sh^{(p)}(G,X)$, then we have either $m = n$ or $p = \mathfrak{p}\overline{\mathfrak{p}}$ with two distinct prime ideals $\mathfrak{p}$ and $\overline{\mathfrak{p}}$ in F.*

Proof. By (unr), we identify $O_{D,p} = M_d(O_p)$ for the integer ring O of F. We write $\varepsilon = \mathrm{diag}[1,0,\dots,0] \in O_{D,p}$. Suppose we have an ordinary triple $(A_0, \overline{\lambda}_0, \eta_0^{(p)})$. By Hilbert's zero theorem (Nullstellensatz; [ALG] I.1.3A), the ordinary locus of $Sh^{(p)}(G,X)$ has $\mathbb{F}$-points $(\mathbb{F} = \overline{\mathbb{F}}_p)$; so, we may assume that A_0 is defined over a finite field $\mathbb{F}_Q$ for $Q = p^r$. Let $(A,\lambda,\eta^{(p)})$ be the abelian variety over $\mathcal{W}$ $(\mathcal{W} = i_p^{-1}(W))$ corresponding to the Serre–Tate coordinate $q_A = 0$ (the canonical lift). Then A has complex multiplication in addition to the multiplication by O_D. We may assume that (A,λ) is defined over a valuation ring $\mathcal{W}_0$ of a number field. Put $W_0 = \varprojlim_j \mathcal{W}_0/p^j\mathcal{W}_0 \cong W(\mathbb{F}_Q) \subset W$ for the finite field $\mathbb{F}_Q$. Write K for the field of fractions of W_0. By the self-duality via λ, the locally free group scheme $A[\mathfrak{p}^j]$ is isomorphic to the Cartier dual of $A[\overline{\mathfrak{p}}^j]$, and $\mathrm{Hom}(A[\mathfrak{p}^j]^{et}_{/\mathcal{W}_0}, \mu_{p^j}) \cong A[\overline{\mathfrak{p}}^j]^{\circ}_{/\mathcal{W}_0}$. Taking the limit, we have

$$\mathrm{Hom}(\mathcal{T}A[\mathfrak{p}^\infty]^{et}, \mathbb{Z}_p(1)) \cong \varprojlim_j A[\overline{\mathfrak{p}}^j]^{\circ}(\overline{\mathbb{Q}}), \tag{8.18}$$

whose right-hand side is written as $\mathcal{T}^{\circ}_{\overline{\mathfrak{p}}}(A)$. From (det), in the same manner as in the proof of Proposition 4.1, we see

$$\varepsilon \cdot Lie(A_{/W_0}) \cong \varepsilon \cdot \mathcal{T}_p^{\circ}(A) \otimes_{\mathbb{Z}_p} W_0 \quad \text{as } W_0 \otimes_{\mathbb{Z}} O\text{-modules}, \tag{8.19}$$

where $\mathcal{T}_p^{\circ}(A) = \varprojlim_j A[p^j]^{\circ}(\overline{\mathbb{Q}})$. By the definition of the signature (m,n), on $\varepsilon \cdot Lie(A_{/K})$, the field F acts through a semi-simple representation whose character is given by $m \cdot \mathrm{id} + n \cdot c$.

Suppose $\mathfrak{p} = \overline{\mathfrak{p}}$ for the moment (so, $\mathfrak{p} = (p)$ under (unr)). Then $\mathcal{T}_p^{\circ}(A)$ gives a maximal isotropic subspace of $\mathcal{T}_p(A) = \varprojlim_j A[p^j](\overline{\mathbb{Q}})$ under the Weil pairing. Since O_p is a valuation ring quadratic over $\mathbb{Z}_p$, if O_p acts on a $\mathbb{Z}_p$-module M of finite rank, its action on $M \otimes_{\mathbb{Z}_p} W_0$ has trace $j(\mathrm{id} + c)$ for a positive integer j. Thus id and complex conjugation c have the same multiplicity on the $\mathbb{Z}_p$-module $\varepsilon \cdot Lie(A_{/K})$, and we have $m = n$.

Now we assume that $\mathfrak{p} \neq \overline{\mathfrak{p}}$ with $\mathfrak{p}$ giving rise to the p-adic place of F corresponding to $i_p \circ \mathrm{id}$. Then by the same reasoning as above, we have

$$\mathcal{T}_p^{\circ}(A) \otimes_{\mathbb{Z}_p} K = \left(\mathcal{T}^{\circ}_{\mathfrak{p}}(A) \oplus \mathcal{T}^{\circ}_{\overline{\mathfrak{p}}}(A)\right) \otimes_{\mathbb{Z}_p} K \cong Lie(A_{/K}),$$

$m = \mathrm{rank}_{O_{\mathfrak{p}}} \varepsilon \cdot \mathcal{T}^{\circ}_{\mathfrak{p}}(A)$ and $n = \mathrm{rank}_{O_{\overline{\mathfrak{p}}}} \varepsilon \cdot \mathcal{T}^{\circ}_{\overline{\mathfrak{p}}}(A)$, where $\mathcal{T}^{\circ}_{\mathfrak{p}}(A) = \varprojlim_j A[\mathfrak{p}^j]^{\circ}(\overline{\mathbb{Q}})$. The existence of p-ordinary A does not impose any restriction on the signature (m,n). □

Remark 8.11 The converse of Lemma 8.10 should be true. Here are some examples of how to create an ordinary couple $(A,\lambda)_{/\mathcal{W}}$ on $Sh^{(p)}(G,X)$ for a fairly general class of data $(D,V,\langle\cdot,\cdot\rangle)$ as in Section 7.1.1.

1. Suppose that $(p)=\mathfrak{p}\overline{\mathfrak{p}}$ with $\mathfrak{p}\neq\overline{\mathfrak{p}}$ in F, $V=F^r$ (row vector space) and $\langle x,y\rangle=\mathrm{Tr}_{F/\mathbb{Q}}(x\delta H{}^t\overline{y})$ for a diagonal matrix $H\in GL_r(\mathbb{Q})$ and $\delta\in F$ with $\overline{\delta}=-\delta$. We further suppose that δ has a positive imaginary part under the identity embedding id of F into $\mathbb{C}$. We take an ordinary elliptic curve E with complex multiplication by O over a discrete valuation ring $\mathcal{W}$ with finite residue field $\mathbb{F}$. We may assume that $L=L^+\oplus L^-$ so that H is positive definite on L^+ and negative definite on L^-. Define $A=(L^+\otimes_{O,\mathrm{id}}E)\times(L^-\otimes_{O,c}E)$, where writing $E(\mathbb{C})=\mathbb{C}/i_\infty(O)$, we regard E as an O-module by sending $\alpha\in O$ to multiplication by $i_\infty(\alpha)$ on $\mathbb{C}/i_\infty(O)$, and $L^-\otimes_{O,c}E$ is the tensor product with L^- twisting the action of O on E by c (i.e., $\alpha\in O$ is sent to multiplication by $i_\infty(c(\alpha))$ on $E(\mathbb{C})=\mathbb{C}/i_\infty(O)$). Then $\mathcal{T}^\circ_{\mathfrak{p}}(A)=(\mathcal{T}^\circ(E)\otimes_{O_{\mathfrak{p}}}L^+)\oplus(\mathcal{T}^\circ(E)\otimes_{O_{\overline{\mathfrak{p}}}}L^-)$. Write H^+ (resp. H^-) for the restriction of H to L^+ and L^-. We have $H=\mathrm{diag}[H^+,H^-]$. Put $|H|=\mathrm{diag}[H^+,-H^-]$ (a positive majorant of H in the sense of Siegel). If we identify $L=(L^+\otimes_{O,\mathrm{id}}O)\oplus(L^-\otimes_{O,c}O)$, $(\ell\otimes\alpha,\ell'\otimes\alpha'):=\mathrm{Tr}_{F/\mathbb{Q}}((\delta\alpha\alpha'^c)\cdot(\ell|H|\cdot{}^t\ell'^c))$ $(\alpha,\alpha'\in O$ and $\ell,\ell'\in L)$ coincides with $\langle\cdot,\cdot\rangle$ on the original L. We take a polarization on E giving an alternating form $(\alpha,\alpha')=\mathrm{Tr}_{F/\mathbb{Q}}(\delta\alpha\alpha'^c)$. Then the tensor product of the polarization on E and $|H|$ on L gives the desired polarization λ which induces the given alternating form $\langle\cdot,\cdot\rangle=\mathrm{Tr}_{F/\mathbb{Q}}(\delta H)$ on $\mathcal{T}(A)=\left(\widehat{L}^+\otimes_{O,\mathrm{id}}O\right)\oplus\left(\widehat{L}^-\otimes_{O,c}O\right)=\widehat{L}$. The triple $(A,\lambda,\eta^{(p)})_{/\mathcal{W}}$, choosing a level structure $\eta^{(p)}$ extending $\mathcal{W}$ suitably, gives a desired ordinary couple on $Sh^{(p)}(GU(m,n),X)(\mathcal{W})$ for the signature (m,n) of H. We can generalize this construction to the case where D is a division algebra and $V=D^r$ in an obvious way.
2. Suppose that D is a division algebra, $V=D^m\oplus D^m$ (row vectors), and $\langle(x_1,x_2),(y_1,y_2)\rangle=\mathrm{Tr}_{D/\mathbb{Q}}(x_1{}^ty_2^*-x_2{}^ty_1^*)$, and $L=L'\oplus L'$ for $L'=O_D^m$ under $\langle\cdot,\cdot\rangle$. Define another pairing $(\ ,\ ):L'\times L'\to\mathbb{Q}$ by $(x,y)=\mathrm{Tr}_{D/\mathbb{Q}}(x^ty^*)$. We take an elliptic curve $E'_{/\mathcal{W}}$ with complex multiplication by the integer ring O' of another imaginary quadratic field F' where p splits (F' can be F if p splits in F). We define $A=E'\otimes_{\mathbb{Z}}L'$. Take the self-dual polarization on E' and identify $H_1(E',\mathbb{Z})$ with $\mathbb{Z}^2$ so that the intersection product is given by $((x_1,x_2),(y_1,y_2))=x_1y_2-x_2y_1$ on $\mathbb{Z}^2$. Then the tensor product $(\ ,\ )'\otimes(\ ,\ )$ on $L=L'\oplus L'=\mathbb{Z}^2\otimes_{\mathbb{Z}}L'=H_1(A,\mathbb{Z})$ gives the original pairing $\langle\cdot,\cdot\rangle$. We take the corresponding polarization λ on $A=E'\otimes_{\mathbb{Z}}L'$. Then (A,λ) satisfies the desired requirements for $U(m,m)$.

Hereafter we assume that

(ord) *$Sh^{(p)}(G,X)(\overline{\mathbb{F}}_p)$ has a point carrying an ordinary abelian scheme.*

Take a closed point $x \in Sh^{(p)}(G,X)(\mathbb{F})$ carrying an ordinary abelian scheme $A_{0/\mathbb{F}}$. We determine the formal completion along x of $Sh^{(p)}(G,X)$ as a formal subscheme of the Serre–Tate deformation space of $(A_0,\lambda_0)_{/\mathbb{F}}$. For an Artinian local W-algebra R with residue field $\mathbb{F} = W/\mathfrak{m}_W = \overline{\mathbb{F}}_p$, we take a deformation $(A,\iota_A,\lambda) \in \widehat{\mathcal{P}}_{A_0,\lambda_0}(R)$ with a level structure $\eta : V(\mathbb{A}^{(p\infty)}) \cong V^{(p)}(A)$ and $i : O_D \hookrightarrow \mathrm{End}(A_{/R})$ giving rise to a point in $Sh^{(p)}(G,X)(R)$. We suppose $\eta^{(p)}(\widehat{L}^{(p)}) = \mathcal{T}^{(p)}(A)$. Via ι_A, we recover the point x given by $(A_0,\lambda_0,i_0,\eta_0^{(p)}) = (A,\lambda,i,\eta^{(p)})\otimes_R \mathbb{F}$ of $Sh^{(p)}(G,X)(\mathbb{F})$. Write $A_\varepsilon := \varepsilon(A[p^\infty])$. Since $A[p^\infty] \cong A_\varepsilon^d$ as Barsotti–Tate p-divisible groups, the abelian scheme A as a deformation of $A_0 = A\otimes_R \mathbb{F}$ is determined by A_ε.

First we assume that $(p) = \mathfrak{p}\overline{\mathfrak{p}}$ with $\mathfrak{p} \neq \overline{\mathfrak{p}}$; so, we have $O_{D,p} = O_D \otimes_{\mathbb{Z}} \mathbb{Z}_p \cong M_d(O_{\mathfrak{p}}) \oplus M_d(O_{\overline{\mathfrak{p}}})$. Identify $V = D^r$. Supposing that $A_{/\mathbb{F}}$ is ordinary, we have $Lie(A_{/\mathbb{F}}) \cong Lie(A[p]^\circ_{/\mathbb{F}})$, where $A[p]^\circ$ is the connected component of $A[p]$. Thus we may assume that $\mathcal{T}A[\mathfrak{p}^\infty]^{et} \cong M_{d\times n}(O_{\mathfrak{p}})^r$ and $\mathcal{T}A[\overline{\mathfrak{p}}^\infty]^{et} \cong M_{d\times m}(O_{\overline{\mathfrak{p}}})^r$. We write the $O_{\mathfrak{p}}$-component of $\mathcal{T}_pA_\varepsilon^{et} = \mathcal{T}_pA_{0,\varepsilon}^{et}$ as $\mathcal{T}_{\mathfrak{p}}A_\varepsilon^{et}$. Then the symmetric pairing $q_{A/R}(\ ,\ ) : \mathcal{T}A_0[p^\infty]^{et} \times \mathcal{T}A_0[p^\infty]^{et} \to \widehat{\mathbb{G}}_m(R)$ induces a homomorphism $q_{A/R} : \mathcal{T}_{\mathfrak{p}}A_\varepsilon^{et} \otimes_{\mathbb{Z}_p} \mathcal{T}_{\overline{\mathfrak{p}}}A_\varepsilon^{et} \to \widehat{\mathbb{G}}_m(R)$, because we have chosen ε with $\varepsilon^* = \varepsilon$, and the pairing is c-Hermitian (i.e., the involution $*$ induces complex conjugation c). The étale level N-structure outside p always lifts uniquely to deformations (because $A[N]_{/R}$ is étale over R), and the lift $i : O_D \hookrightarrow \mathrm{End}(A_{/R})$ of i_0 is unique if it exists, because $\iota : O_D \hookrightarrow \mathrm{End}(A_{/R})$ is determined by $\iota : O_D \hookrightarrow \mathrm{End}(A[\ell^\infty]_{/R})$ for a prime $\ell \neq p$ and hence by $\eta_0^{(p)}$. Thus we can ignore the level structure while we study deformations of $(A_0,\lambda_0,i_0,\eta_0^{(p)})_{/\mathbb{F}}$. So we consider the functor

$$\widehat{\mathcal{P}}_{A_0,\lambda_0,\imath_0,\iota_0}(R) = \left[(A_{/R},i,\iota_A,\lambda)\middle|(A,\iota_A,\lambda) \in \widehat{\mathcal{P}}_{A_0,\lambda_0}(R) \text{ and } \iota_A \text{ induces } \iota_0\right].$$

Then the above argument combined with Theorem 8.9 shows

$$\widehat{\mathcal{P}}_{A_0,\lambda_0,\imath_0,\iota_0}(R) \cong \mathrm{Hom}_{\mathbb{Z}_p}(\mathcal{T}_{\mathfrak{p}}A_\varepsilon^{et} \otimes_{\mathbb{Z}_p} \mathcal{T}_{\overline{\mathfrak{p}}}A_\varepsilon^{et}, \widehat{\mathbb{G}}_m(R)) =: \widehat{S}(R), \tag{8.20}$$

because the c-Hermitian form on $(\mathcal{T}_{\mathfrak{p}}A_\varepsilon^{et} \times \mathcal{T}_{\overline{\mathfrak{p}}}A_\varepsilon^{et}) \otimes (\mathcal{T}_{\mathfrak{p}}A_\varepsilon^{et} \times \mathcal{T}_{\overline{\mathfrak{p}}}A_\varepsilon^{et})$ is determined by its restriction on $(\mathcal{T}_{\mathfrak{p}}A_\varepsilon^{et} \times \{0\}) \times (\{0\} \times \mathcal{T}_{\overline{\mathfrak{p}}}A_\varepsilon^{et})$.

We now assume that p is inert in F; so, $m = n$. For simplicity, we assume that $D = F$, $V = F^{2m}$, and $\langle (x_1,x_2),(y_1,y_2)\rangle = \mathrm{Tr}_{F/\mathbb{Q}}(x_1\,{}^t\overline{y}_2 - x_2\,{}^t\overline{y}_1)$. We take $L_p = O_p^m \oplus O_p^m$ and write the first factor O_p^m as $L_{1,p}$. Then $C = M_{2m}(F)$ and $x^\iota = J_m\,{}^t\overline{x}J_m^{-1}$ for $x \in M_{2m}(F)$. Then by the same computation as above

$$\widehat{\mathcal{P}}_{A_0,\lambda_0,\imath_0,\iota_0}(R) \cong \mathrm{Hom}_{\mathbb{Z}_p}(\mathcal{T}_pA_\varepsilon^{et} \otimes_{O_p} \mathcal{T}_pA_\varepsilon^{et}, \widehat{\mathbb{G}}_m(R)) = \widehat{S}(R), \tag{8.21}$$

and $\mathcal{T}_pA_\varepsilon^{et} \otimes_{O_p} \mathcal{T}_pA_\varepsilon^{et} \cong M_m(O_p)$, identifying $L_{1,p}$ with $\mathcal{T}_pA_\varepsilon^{et}$.

The case where $F = \mathbb{Q}$ (so the involution is of the first kind) is similar to the above inert case. We assume that G is of type C. Since $G(\mathbb{Q}_p) \cong$

$GSp_{2n}(\mathbb{Q}_p)$ (i.e., we may take $s = J_n$ and $* = \mathrm{id}$ by the argument above Remark 7.4 and (7.6)), similarly to the symplectic case, we have

$$\widehat{\mathcal{P}}_{A_0,\lambda_0,\iota_0,\iota_0}(R) \cong \mathrm{Hom}_{\mathbb{Z}_p}(Sym^2(\mathcal{T}_p A_\varepsilon^{et}), \widehat{\mathbb{G}}_m(R)) = \widehat{S}(R). \tag{8.22}$$

The results as above (8.20) through (8.22) show the liftability of an ordinary point of $Sh^{(p)}(G,X)(\mathbb{F})$ to characteristic 0 (as used in the proof of Theorem 7.1).

8.3 Vertical Control Theorem

We prove the VCT for projective Shimura varieties of PEL type. So we return to the notation introduced in Section 7.1. For simplicity, we give a full exposition under the following four assumptions.

(u1) $F = \mathbb{Q}[\sqrt{-\delta}]$ is an imaginary quadratic field with discriminant $-\delta$;
(u2) $D_{/F} \neq F$ is a division algebra with positive involution of the second kind;
(u3) $V = D$; so, $GU(\mathbb{Q}) \subset C^\times$ for the opposite algebra $C = {}^tD$ of D;
(u4) The prime p splits into two distinct primes $\mathfrak{p}$ and $\overline{\mathfrak{p}}$ in O_F.

We write $O = O_F$ in this section.

Under (u1), we have $G = GU$; so, we have an advantage that any statements already given exclusively for G or GU are both valid for our $G = GU$. Since $F/\mathbb{Q}$ is a Galois extension, the reflex field $E = E(G,X)$ is a subfield of F; so, $O_{E,\mathfrak{p}} = \mathbb{Z}_p$ by (u4). By (u3), $Sh^{(p)}(G,X)$ is projective over $\mathcal{V}$. To have the ordinary locus, the condition (ord) in the previous section is always in force (which can be easily verified as in Remark 8.11 in many cases). It is not difficult to generalize our result to more general projective Shimura varieties of PEL type, although the endeavor certainly complicates the notation (see [H02] for general treatment including quasi-projective Shimura varieties). At the end of this section, we touch briefly upon the case where p is inert. By our simplifying assumption, the reflex field for (GU,X) is either $\mathbb{Q}$ or F.

Recall some of the notation introduced in Section Section 7.1 in our special case. As in 7.1.1, let $G = GU$ be a unitary similitude group defined with respect to a central division algebra D over F with a positive involution $*$. Fix a prime p, and take $V = D$ and a left O_D-ideal L with $L \otimes_{\mathbb{Z}} \mathbb{Z}_p = O_D \otimes_{\mathbb{Z}} \mathbb{Z}_p$. Pick an element with $s \in D^\times$ with $s^* = -s$, and define an alternating form $\langle x,y\rangle = \mathrm{Tr}_{D/\mathbb{Q}}(xsy^*)$ for $x,y \in V$, where $\mathrm{Tr}_D : D \to \mathbb{Q}$ is the reduced trace composed with $\mathrm{Tr}_{F/\mathbb{Q}} : F \to \mathbb{Q}$. Then we see $\langle bx,y\rangle = \mathrm{Tr}_{D/\mathbb{Q}}(bxsy^*) = \mathrm{Tr}_{D/\mathbb{Q}}(xsy^*b) = \mathrm{Tr}_{D/\mathbb{Q}}(xs(b^*y)^*) = \langle x, b^*y\rangle$ for $x,y \in V$ and $b \in D$.

Suppose that L satisfies (L1–2) in Section 7.1.1 for this alternating form; so, $s \in O_{D,p}^\times$. As before, we write $\iota : C \to C$ for the involution on C given by $\langle cx,y\rangle = \langle x, c^\iota y\rangle$; then, $c^\iota = sc^*s^{-1}$. Recall

$$GU(\mathbb{Q}) = G(\mathbb{Q}) = \left\{\gamma \in C^\times \middle| \gamma\gamma^\iota \in F_0 = \mathbb{Q}\right\},$$
$$U(\mathbb{Q}) = \left\{\gamma \in GU(\mathbb{Q}) \middle| \gamma\gamma^\iota = 1\right\} \text{ and } G_1(\mathbb{Q}) = \left\{\gamma \in U(\mathbb{Q}) \middle| N(\gamma) = 1\right\},$$

where $N : C \to F$ is the reduced norm map. Suppose (unr) for D ((sc) for G_1 is automatic).

8.3.1 Hecke Operators on Deformation Space

In order to state the control theorem, we need to give a precise definition of the projector e_G, and we devote this subsection to proving the integrality of suitable p-Hecke operators (which yield e_G after taking a p-adic limit). Let O_C be the order of L in C: $O_C = \{x \in C | xL \subset L\}$. We write $U(\widehat{\mathbb{Z}})$ for $\widehat{O}_C^\times \cap U(\mathbb{A}^{(\infty)})$, where $\widehat{O}_C = O_C \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}$. We fix an open compact subgroup $K \subset U(\widehat{\mathbb{Z}})$ such that $K = K_p \times K^{(p)}$ with $K_p \cong GL_{m+n}(\mathbb{Z}_p)$ via the projection to the $\mathfrak{p}$-factor.

We consider p-ordinary test objects $\underline{A} = (A, \lambda, i, \overline{\eta}^{(p)})_{/R}$ over a local Artinian W-algebra R. Since the pairing $q_{A/R} \in \widehat{S}(R)$ is actually determined by its restriction to $q_{A/R} : \mathcal{T}_{\mathfrak{p}} A_\varepsilon^{et} \times \mathcal{T}_{\overline{\mathfrak{p}}} A_\varepsilon^{et}$, we look into this restriction. We study the $O_{\mathfrak{p}}$-linear endomorphism algebra $\mathrm{End}_{BT}(\varepsilon \cdot A[\mathfrak{p}^\infty]_{/R})$ of the Barsotti–Tate group $\varepsilon \cdot A[\mathfrak{p}^\infty]_{/R}$. Write each endomorphism α as $\left(\begin{smallmatrix} a_\alpha & b_\alpha \\ 0 & d_\alpha \end{smallmatrix}\right)$ with

$$a_\alpha \in \mathrm{End}_{BT}(\varepsilon \cdot A[\mathfrak{p}^\infty]^\circ),\ b_\alpha \in \mathrm{Hom}_{BT}(\varepsilon \cdot A[\mathfrak{p}^\infty]^{et}, \varepsilon \cdot A[\mathfrak{p}^\infty]^\circ) \text{ and}$$
$$d_\alpha \in \mathrm{End}_{BT}(\varepsilon \cdot \mathcal{T}_{\mathfrak{p}}^{et}(A)).$$

If R is an algebraically closed field of characteristic p, the étale-connected exact sequence $\varepsilon \cdot A_0[\mathfrak{p}^\infty]^\circ \hookrightarrow \varepsilon \cdot A_0[\mathfrak{p}^\infty] \twoheadrightarrow \varepsilon \cdot A_0[\mathfrak{p}^\infty]^{et}$ is (noncanonically) split. In any case, α acts on $\mathcal{T}_{A,\mathfrak{p}} = \mathcal{T}_{\mathfrak{p}} A_\varepsilon^{et} \oplus \mathcal{T}_{\overline{\mathfrak{p}}} A_\varepsilon^{et}$ diagonally via a_α and d_α. We regard $T = \mathbb{G}_m^n$ as a maximal split torus of $GL_m(O_{\overline{\mathfrak{p}}}) \times GL_n(O_{\mathfrak{p}})$, which is the automorphism group of $\varepsilon \cdot A[\mathfrak{p}^\infty]^{et} \times \varepsilon \cdot A[(\mathfrak{p}^c)^\infty]^{et}$.

Let $\mathbb{X}_{/Sh_K^{(p)}}$ be the universal abelian scheme for an open compact subgroup $K \subset G(\mathbb{A}^{(\infty)})$ maximal at p. We write, for integers k and k', $P_k = \varepsilon \cdot (\mathbb{X}_{/S}[\mathfrak{p}^k]^{et})$, $P_\infty = \bigcup_k P_k$, $P'_{k'} = \varepsilon \cdot (\mathbb{X}_{/S}[\overline{\mathfrak{p}}^{k'}]^{et})$, and $P'_\infty = \bigcup_{k'} P'_{k'}$. We apply the theory developed in Section 8.1 to each piece P_k and $P'_{k'}$; so, we obtain the theory of false modular forms for $GL_n(O_{\mathfrak{p}}) \times GL_m(O_{\overline{\mathfrak{p}}})$ ($O_{\mathfrak{p}} \cong O_{\overline{\mathfrak{p}}} \cong \mathbb{Z}_p$). Since p is unramified in F, we have $O \otimes_{\mathbb{Z}_p} W \cong W^I$ for the set of embedding $I = \{\sigma = \mathrm{id}, c\}$ of F into $\overline{\mathbb{Q}}$. Then we consider filtrations fil_σ and fil_c of $\varepsilon(A[\mathfrak{p}^k]^{et})$ and $\varepsilon(A[\overline{\mathfrak{p}}^{k'}]^{et})$ and test objects of the following type: $\underline{A}_{(k,k')} = \{A, \lambda, i, \overline{\eta}^{(p)}, fil_\sigma, fil_c\}_{/R}$. Let $M = Sh_K^{(p)}$ and $S = M\left[\frac{1}{E}\right]$, where E is a lift of the Hasse invariant. We write $T^B_{\ell,k,k'/S_\ell}$ for the étale covering over $S_\ell = S \otimes_W W_\ell$ ($W_\ell = W/p^\ell W$) classifying the above test objects. Similarly, $T_{\ell,k,k'}$ classifies $(A, \lambda, i, \overline{\eta}^{(p)}, (O/\mathfrak{p}^k)^n \times (O/\overline{\mathfrak{p}}^{k'})^m \cong \varepsilon \cdot A[\mathfrak{p}^k]^{et} \times \varepsilon \cdot A[\overline{\mathfrak{p}}^{k'}]^{et})$. The covering $T_{\ell,\infty,\infty}/S_\ell$ is an étale faithful covering with Galois group canonically containing $M_p = GL_m(O_{\overline{\mathfrak{p}}}) \times GL_n(O_{\mathfrak{p}})$, and we have $T_{\ell,\infty,\infty}/M_p \cong S_\ell$.

We have an action of the expanding semi-group on filtrations of P_k for the étale sheaf P_k as described in Section 8.1.6 and on the abelian scheme $\mathbb{X}$ through isogenies (this isogeny action on abelian schemes A was not specified when we described the action on filtration in Section 8.1.6). If we have a p-isogeny $\beta : P_\infty \to P_\infty$ preserving a filtration of P_∞, we may assume that the matrix form of β is given by $\beta_\jmath = \begin{pmatrix} 1_{n-\jmath} & * \\ 0 & p1_\jmath \end{pmatrix}$ with respect to a base compatible with the filtration. Then the action of $\beta_\jmath$ is to give a new filtration on P_∞.

Since we cannot separate P_k and $P'_{k'}$ (which is sitting in the single universal abelian scheme $\mathbb{X}$), we define $\alpha_{m+\jmath} = \mathrm{diag}[\beta_\jmath, p1_m]$ and consider an isogeny of type $\alpha_{m+\jmath}$. We can thus interpret the operator action in terms of the sextuple $(A, \lambda, i, \overline{\eta}^{(p)}, fil_\sigma, fil_c)_{/\mathbb{F}}$ as follows. Take an isogeny $\alpha : A \to A_\alpha$ of type $\alpha_{m+\jmath}$ as above (inducing α on $P_\infty \oplus P'_\infty$). Then we get a new filtration $\alpha(fil_\sigma, fil_c)$. The p-isogeny is insensitive to the level K-structure (for K maximal at p), and A_α has an induced polarization, still written as λ. Then we have

$$f|(t_\jmath(\mathfrak{p}) \times t_m(\overline{\mathfrak{p}}))(A, \lambda, i, \overline{\eta}^{(p)}, fil_\sigma, fil_c) = \sum_\alpha \widetilde{\rho}_\kappa(\alpha^{-1}) f(A_\alpha, \lambda, i, \overline{\eta}^{(p)}, \alpha(fil_\sigma, fil_c)). \quad (8.23)$$

Thus we have a $GL(n) \times GL(m)$-Hecke operator $t_\jmath(\mathfrak{p}) \times t_m(\overline{\mathfrak{p}})$ acting on the coherent cohomology $H^0(S_\ell, \underline{\omega}^\kappa)$. This is actually an *over-simplified* version. The exact sequence $\mathrm{Hom}(P'_k, \mathbb{G}_m) \hookrightarrow \mathbb{X}[\mathfrak{p}^k] \twoheadrightarrow P_k$ may not split over S_ℓ; so, the isogeny α can be defined only over a finite flat extension $S_\ell^{\alpha_{m+\jmath}}$ of S_ℓ (which could be inseparable over an étale extension of S_ℓ). In other words, if we replace the sum $\sum_\alpha \widetilde{\rho}_\kappa(\alpha^{-1}) f(A_\alpha, \lambda, i, \overline{\eta}^{(p)}, \alpha(fil_\sigma, fil_c))$ in (8.23) by the trace $\mathrm{Tr}_{S_\ell^{\alpha_{m+\jmath}}/S_\ell}(\widetilde{\rho}_\kappa(\alpha^{-1}) f(A_\alpha, \lambda, i, \overline{\eta}^{(p)}, \alpha(fil_\sigma, fil_c)))$ as in (4.18) of Section 4.1.10, we can relate $t_\jmath$ to a global Hecke operator $U_\imath(\mathfrak{p})$ which is divisible by the degree of $S_\ell^{\alpha_{m+\jmath}}$ over the maximal étale cover of S_ℓ under $S_\ell^{\alpha_{m+\jmath}}$.

For the moment, we pretend that the over-simplified version (8.23) is valid (and we later justify our argument). Thus for a while, our argument is heuristic. Here we resort to the theory described in Section 8.1.7 applied to $GL_m \times GL_n$, although the construction of the idempotent $e = e_G$ is global (so some modification is necessary). So, under the notation used in Section 8.1.7, we relate e_G to $e_{GL_n} \times e_{GL_m}$.

Since $U(\mathbb{Q}_p) \cong GL_{m+n}(\mathbb{Q}_p)$, we embed $GL(n) \times GL(m)$ into $U(\mathbb{Q}_p)$ by $(x, y) \mapsto \mathrm{diag}[x, y]$. This implies the p-isogeny whose kernel sits in $A[\mathfrak{p}^\infty]^{et}$ (resp. $A[\overline{\mathfrak{p}}^\infty]^{et}$) corresponds to x (resp. y). Then write $P = P_{m,n}$ for the upper triangular parabolic subgroup of U whose Levi subgroup is given by the image of $GL(n) \times GL(m)$. Let $N_{m,n}$ be the unipotent radical of $P_{m,n}$.

Write $\alpha_\jmath \in GU(\mathbb{Q}_p)$ whose projection to $C_\mathfrak{p} = C \otimes_F F_\mathfrak{p}$ is given by $\mathrm{diag}[1_{m+n-\jmath}, p \cdot 1_\jmath]$ and $\nu(\alpha_\jmath) = p$. We then have Hecke operators

$$U_j(\mathfrak{p}) = U(\alpha_\jmath) = N_G \alpha_\jmath N_G,$$

where N_G is the upper unipotent subgroup of $G(\mathbb{Z}_p)$. Thus $N = N_G/N_{m,n}$ and $GL(m)\times GL(n) = P_{m,n}/N_{m,n}$. Since we identify $U(\mathbb{Q}_p)$ with $GL_{m+n}(\mathbb{Q}_p)$ by projecting down $C \otimes_{\mathbb{Q}} \mathbb{Q}_p$ to the first component $C_{\mathfrak{p}}$, as a double coset, we see (symbolically)

$$U_j(\mathfrak{p}) = \begin{cases} \bigsqcup_{u\in N_{m,n}/\alpha_j^{-1}N_{m,n}\alpha_j} N_G(t_{j-n}(\mathfrak{p}) \times t_m(\overline{\mathfrak{p}}))u & \text{if } j > n, \\ \bigsqcup_{u\in N_{m,n}/\alpha_j^{-1}N_{m,n}\alpha_j} N_G(t_0(\mathfrak{p}) \times t_j(\overline{\mathfrak{p}}))u & \text{if } j \le n, \end{cases}$$

where we mean, for example, by $(t_{j-n}(\mathfrak{p}) \times t_m(\overline{\mathfrak{p}}))$, the double coset

$$N(\mathrm{diag}[1_{2n-j}, p\cdot 1_{j-n}] \times p\cdot 1_m)N$$

in $GL_n(F_{\mathfrak{p}}) \times GL_m(F_{\overline{\mathfrak{p}}})$ for the upper triangular unipotent subgroup $N \subset GL_n(O_{\mathfrak{p}}) \times GL_m(O_{\overline{\mathfrak{p}}})$. This shows that the Hecke operator $U_j(\mathfrak{p})$ induces

$$[N_{m,n} : \alpha_j^{-1}N_{m,n}\alpha_j](t_{j-n}(\mathfrak{p}) \times t_m(\overline{\mathfrak{p}})) \quad \text{or} \quad [N_{m,n} : \alpha_j^{-1}N_{m,n}\alpha_j](t_0(\mathfrak{p}) \times t_m(\overline{\mathfrak{p}}))$$

accordingly as $j > n$ or not. By computation, we get the following heuristic multiplicity formula for the weight $\mu_{m,n} \in X(T)$ as in (5.38):

$$[N_{m,n} : \alpha_j^{-1}N_{m,n}\alpha_j] = \mu_{m,n}(\alpha_j) = \begin{cases} |p|_{\mathfrak{p}}^{-n(m+n-j)} & \text{if } j > n, \\ |p|_{\mathfrak{p}}^{-mj} & \text{if } j \le n. \end{cases} \tag{8.24}$$

This formula suggests us that $U_j(\mathfrak{p})$ is divisible by $\mu_{m,n}(\alpha_j)$, which we justify later.

Since the universal deformation space of $(A, \lambda, i, \overline{\eta}^{(p)}, fil_{\mathfrak{p}}, fil_{\overline{\mathfrak{p}}})_{/\mathbb{F}}$ is isomorphic to $\widehat{S} = \mathrm{Hom}(\mathcal{T}_{\mathfrak{p}}A_\varepsilon^{et} \otimes_{\mathbb{Z}_p} \mathcal{T}_{\overline{\mathfrak{p}}}A_\varepsilon^{et}, \widehat{\mathbb{G}}_m)$, as already seen in Section 8.2.5, we can analyze the effect of an isogeny $\alpha : \mathbf{A}_{/\widehat{S}} \to \mathbf{A}'_{/\widehat{S}}$ of type α_j on $\widehat{S}$ of the universal deformation $\mathbf{A}_{/\widehat{S}}$, which sends

$$\mathrm{Hom}(\mathcal{T}_{\mathfrak{p}}A_\varepsilon^{et} \times \mathcal{T}_{\overline{\mathfrak{p}}}A_\varepsilon^{et}, \widehat{\mathbb{G}}_m) \ni q(x,y) \mapsto \sqrt[p]{q(\alpha(x), \alpha(y))}.$$

Here we have formulated the group structure of $\widehat{\mathbb{G}}_m$ multiplicatively (contrary to the additive notation so far used; so, the right-hand side becomes $q(\alpha(x), \alpha(y))/p$ if we use additive notation). We need to extract a pth root as above by the following reasoning. The coordinate $q \in \widehat{S}$ measures the depth of nonsplitting of the exact sequence $\mathrm{Hom}(P'_k, \mathbb{G}_m) \hookrightarrow \varepsilon(\mathbb{X}[\mathfrak{p}^k]) \twoheadrightarrow P_k$, and the sequence for $k = 1$ is split if q is a p-power. Thus the isogeny α exists over $\widehat{S}^{1/p}$. The isogeny is defined over a smaller covering $\widehat{S}[(q\circ\alpha)^{1/p}] = Spf(\widehat{\mathcal{O}}_S[(q\circ\alpha)^{1/p}])$ by definition; so, $\widehat{S}^\alpha/\widehat{S}$ is given by $\widehat{S}[(q\circ\alpha)^{1/p}]$. At this point, we are taking pth roots, and hence *pure inseparability* (we pretended not to have) comes in. Then the action of the isogeny α of type α_j on $\widehat{S}$ only depends on its effect on $\mathcal{T}_{\mathfrak{p}}A_\varepsilon^{et}$ and $\mathcal{T}_{\overline{\mathfrak{p}}}A_\varepsilon^{et}$, not on the individual choice α. This means that the covering $\widehat{S}^\alpha$ over $\widehat{S}$ carrying the isogeny α only depends on

the image of α in the Levi quotient of P. Indeed, taking a base $(x_\imath)_\imath$ of $\mathcal{T}_{\mathfrak{p}}A_\varepsilon^{et}$ and $(y_k)_k$ of $\mathcal{T}_{\overline{\mathfrak{p}}}A_\varepsilon^{et}$ so that the matrix of the isogeny is exactly $\alpha_\jmath$, the effect on $Q=(Q_{k,l})=(\log(e(x_k,y_l)))$ is given by

$$\left(\begin{smallmatrix}1_m & Q\\ 0 & 1_n\end{smallmatrix}\right)\mapsto \alpha_\jmath\left(\begin{smallmatrix}1_m & Q\\ 0 & 1_n\end{smallmatrix}\right)\alpha_\jmath^{-1}, \tag{8.25}$$

and S_1^α has the degree of pure inseparability given by the value in (8.24).

Here is the justification of our argument. Write the multiplicative variable on $\widehat{S}$ as an $m\times n$ variable matrix $t=(t_{k,l})$ (so $Q_{k,l}=\log(t_{k,l})$). The conjugation $\left(\begin{smallmatrix}1_m & Q\\ 0 & 1_n\end{smallmatrix}\right)\mapsto \alpha_\jmath\left(\begin{smallmatrix}1_m & Q\\ 0 & 1_n\end{smallmatrix}\right)\alpha_\jmath^{-1}$ induces $Q_{k,l}\mapsto p^{-1}Q_{k,l}$ for some indices (k,l). We split the set of indices (k,l) into a disjoint union $I\sqcup J$ of two subsets so that the conjugation by $\alpha_\jmath$ induces $Q_{k,l}\mapsto p^{-1}Q_{k,l}$ if and only if $(k,l)\in I$. The covering $\widehat{S}_\ell^{\alpha_\jmath}$ is given by $Spf(\widehat{\mathcal{O}}_S[t_{k,l}^{1/p}]_{(k,l)\in I})$. Thus a formal function on $\widehat{S}$ has expansion $\sum_\xi a_\xi t_\xi\in \widehat{W[t_{k,l},t_{k,l}^{-1}]}$ for $\xi\in\mathbb{Z}^{I\sqcup J}$. Writing $\xi(I)$ for the I-part of the index ξ, a formal function f on $\widehat{S}_\ell^{\alpha_\jmath}$ has expansion $f=\sum_\xi a_\xi t^{\xi(I)/p+\xi(J)}$, and we have

$$\mathrm{Tr}_{\widehat{S}^{\alpha_\jmath}/\widehat{S}}(f)=\mu_{m,n}(\alpha_\jmath)\sum_{\xi:\xi(I)\equiv 0 \mod p} a_\xi t^{\xi(I)/p+\xi(J)},$$

because $\mu_{m,n}(\alpha_j)$ is the degree of the purely inseparable (wildly ramified) covering $\widehat{S}^{\alpha_\jmath}\twoheadrightarrow\widehat{S}$ and $\mathrm{Tr}(t_{k,l}^{\imath/p})=\mu_{m,n}(\alpha_\jmath)t_{k,l}^{\imath/p}$ or 0 accordingly as $p|i$ or not. Replacing the term $\sum_\alpha \widetilde{\rho}_\kappa(\alpha^{-1})f(A_\alpha,\lambda,i,\overline{\eta}^{(p)},\alpha(fil_\sigma,fil_c))$ in (8.23) by the trace $\mathrm{Tr}_{S^\alpha/S}(\widetilde{\rho}_\kappa(\alpha^{-1})f(A_\alpha,\lambda,i,\overline{\eta}^{(p)},\alpha(fil_\sigma,fil_c)))$, we get the p-divisibility of the operator $U_\jmath(\mathfrak{p})$ as the (heuristic) multiplicity formula (8.24) suggests. This justifies the heuristic argument we gave (the heuristic argument is actually valid for ℓ-adic étale cohomology with $\ell\neq p$ as already explained).

Let $\widehat{S}_\ell=\widehat{S}\otimes_W W_\ell$. On the universal deformation $\mathbf{A}_{/\widehat{S}}$, the sheaf $\varepsilon(Lie(\mathbf{A}))$ over $\widehat{S}_\ell$ is given by $\varepsilon(Lie(\mathbf{A}[p^\ell]^\circ))$. By duality, $\underline{\omega}_{\widehat{S}_\ell}=\mathcal{O}_{\widehat{S}_\ell}\otimes_{\mathbb{Z}_p}\mathbf{A}[p^\ell]^{et}$, which again only depends on $A_0[p^\ell]^{et}$; so, the Hecke operator $U_\jmath(\mathfrak{p})$ is still divisible by $\mu_{m,n}(\alpha_\jmath)=[N_{m,n}:\alpha_\jmath^{-1}N_{m,n}\alpha_\jmath]$ on $\underline{\omega}^\kappa_{\widehat{S}}$ for all $\kappa>0$. Thus the action of the correspondence of characteristic 0 on $H^0(S_\infty,\underline{\omega}^\kappa)$ is exactly a multiple by the number in (8.24) of the operator induced by the mod p correspondence, which is an integral operator. From this our claim is clear.

In any case, we can divide the action of $U_\jmath(\mathfrak{p})$ by the number in (8.24) keeping the integrality of the operator on $\underline{\omega}^\kappa$.

Lemma 8.12 *Let the notation be as above. We have a well-defined integral operator $\mu_{m,n}(\alpha_\jmath)^{-1}U_\jmath(\mathfrak{p})$ on $H^0(T_{m,n},\underline{\omega}^\kappa\otimes\Omega_{S/W})$.*

We then define

$$e_G=\lim_{n\to\infty}(U(\mathfrak{p}))^{n!}, \tag{8.26}$$

where $U(\mathfrak{p})=\prod_{\jmath=1}^{m+n}\mu_{m,n}(\alpha_\jmath)^{-1}U_\jmath(\mathfrak{p})$ for $\mu_{m,n}\in X(T)$ as in (5.38).

The divisibility of $T(\xi)$ by $\mu(\xi)$ for μ in (5.40) for the symplectic group $GSp(2n)$ in the correspondence ring of $Sh^{(p)}(GSp(2n), X)_{/\mathbb{F}_p}$ has been proven in a different manner in [DAV] VII.4 (and also there is another proof of Lemma 8.12 via the q-expansion principle in this case: [H02]).

As before, write $M_\ell = Sh^{(p)}_{K/W} \otimes_W W_\ell$ for $\ell = 1, 2, \ldots, \infty$. As for $T_j(\mathfrak{p})$, if $\kappa \geq \mu_{m,n}$, $\mu_{m,n}(\alpha_j)^{-1} T_j(\mathfrak{p})$ is integral, $\mu_{m,n}(\alpha_j)^{-1} T_j(\mathfrak{p}) \equiv \mu_{m,n}(\alpha_j)^{-1} U_j(\mathfrak{p})$ mod p, and we have the well-defined projector

$$e^\circ_G = \lim_{n \to \infty} (T(\mathfrak{p}))^{n!}, \tag{8.27}$$

where $T(\mathfrak{p}) = \prod_{j=1}^{m+n} \mu_{m,n}(\alpha_j)^{-1} T_j(\mathfrak{p})$.

The proof of the p-integrality of the operator $T(\mathfrak{p})$ in (8.27) and the fact that e_G induces an isomorphism of $H^0_{ord}(M_{/\mathbb{Z}_p}, \underline{\omega}^\kappa)$ onto $H^0_{ord}(S_{\infty/\mathbb{Z}_p}, \underline{\omega}^\kappa)$ if $\kappa \gg \mu_{m,n}$ is similar to the treatment in the Hilbert modular case (see the proofs of Theorems 4.9 and 4.37), and the argument is also analogous to the proof of Theorem 5.3. Indeed, $f|T_j(\mathfrak{p})$ is the sum of $f|U_j(\mathfrak{p})$ and the terms $\widetilde{\rho}_\kappa(\beta^{-1}) f(A_\beta, \cdots)$ for nonétale isogenies $\beta : A \to A_\beta$. The abelian variety A_β is a quotient of A by a finite flat group scheme (so well-defined over W; see [ABV] Section 12). Thus $f(A_\beta, \cdots)$ is well-defined over W. However, $\mu_{m,n}(\alpha_j)^{-1} \widetilde{\rho}_\kappa(\beta^{-1}) \equiv 0 \mod p$ (by the argument proving Theorem 5.3), and hence we get the congruence $\mu_{m,n}(\alpha_j)^{-1} T_j(\mathfrak{p}) \equiv \mu_{m,n}(\alpha_j)^{-1} U_j(\mathfrak{p}) \mod p$. The boundedness of the rank of the topological Betti cohomology (independent of κ) can be proven in the same way as in the Hilbert modular case (see [H95] Theorem 1.1 for a proof). This shows the condition (F) in Section 8.1.7 and hence the isomorphism $H^0_{ord}(M_{/\mathbb{Z}_p}, \underline{\omega}^\kappa) \cong H^0_{ord}(S_{\infty/\mathbb{Z}_p}, \underline{\omega}^\kappa)$ for $\kappa \gg \mu_{m,n}$.

There is an alternative argument showing the p-integrality of the operator $T_j(\mathfrak{p})$ on $\underline{\omega}^\kappa$ over M. Indeed, the well-definedness of $T_j(\mathfrak{p})$ as a correspondence on $Sh^{(p)}_K \times_W Sh^{(p)}_K$ can be seen in the same manner as in the proof of Theorem 6.28 where the Siegel modular case is considered. Once we have a correspondence well-defined, we can (in principle) show the integrality in the case of the Shimura variety of PEL type (see [DAV] Chapters VI and VII for more details in the Siegel modular case).

8.3.2 Statements and Proof

We keep the notation introduced in the previous section. Let N be the upper unipotent subgroup of $GL_m(O_{\overline{\mathfrak{p}}}) \times GL_n(O_{\mathfrak{p}})$. Applying the theory expounded in Section 8.1 to the present setting, we can define the space of p-adic modular forms V on S_∞ (which is the formal completion of S along S_1). We simply write V for V^N in this subsection. Thus $V = \varprojlim_\ell \left(\varinjlim_{k,k'} V^N_{\ell,k,k'} \right)$ for $V_{\ell,k,k'} = H^0(T_{\ell,k,k'}, \mathcal{O}_{T_{\ell,k,k'}})^N$. We also define $\mathcal{V} = \varinjlim_\ell \left(\varinjlim_{k,k'} V^N_{\ell,k,k'} \right)$. We now state the final result on vertical control in this book.

Theorem 8.13 *Let the notation be as above. Let* $(GU, V, D, F, p = \mathfrak{p}\overline{\mathfrak{p}})$ *be as in* (u1–4) *at the beginning of this section. Identify* $\mathbb{Z}_p = O_{\mathfrak{p}}$ *naturally. Suppose that* $M_{/\mathbb{Z}_p} = Sh_K^{(p)}(GU, X)_{/\mathbb{Z}_p}$ *for an open compact subgroup* $K \subset GU(\mathbb{A}^{(\infty)})$ *maximal at* p *is proper smooth over* $\mathbb{Z}_p$ *and that* $S_{/\mathbb{Z}_p} = M\left[\frac{1}{E}\right]$ *for a lift* E *of the Hasse invariant. Let* T *be the diagonal torus of* $GL_m(O_{\overline{\mathfrak{p}}}) \times GL_n(O_{\mathfrak{p}})$ *and* N *be the upper unipotent radical of* $GL_m(O_{\overline{\mathfrak{p}}}) \times GL_n(O_{\mathfrak{p}})$. *Regard* T *as a diagonal torus of the unitary group* $U(\mathbb{Q}_p)$. *We say* $\kappa \in X(T)$ *is positive if* κ *is dominant with respect to the upper triangular Borel subgroup of* $U_{/\mathbb{Z}_p} = GL(m+n)$ *in the sense described in* Section 8.1.2. *We write* $X_+(T)$ *for the set of positive weights in* $X(T)$.

(1) *There exists a canonical inclusion for* $R = \mathbb{Z}_p$ *and* $\mathbb{Q}_p/\mathbb{Z}_p$,

$$\beta : \bigoplus_{\kappa \in X_+(T)} H^0(M_{/\mathbb{Z}_p}, \underline{\omega}^{\kappa} \otimes_{\mathbb{Z}_p} R) \hookrightarrow V \otimes_{\mathbb{Z}_p} R.$$

(2) $\mathrm{Im}(\beta)[\frac{1}{p}] \cap V$ *is dense in* V;

(3) *Write* $U(\mathfrak{p}) = \prod_{1 \le j \le m+n} u_j(\mathfrak{p})$ *for the standard Hecke operators at* p *of level* p^{∞}; *that is,* $u_j(\mathfrak{p}) = \mu_{m,n}(\alpha_j)^{-1} U_j(\mathfrak{p})$ *for the character* $\mu_{m,n} \in X(T)$ *as in* (5.38), *and define the ordinary projector* $e_G = \lim_{n \to \infty} U(\mathfrak{p})^{n!}$ *on* V. *Then* $H^0_{ord}(S, \underline{\omega}^{\kappa} \otimes_{\mathbb{Z}_p} \mathbb{Q}_p/\mathbb{Z}_p)$ *is a* p*-divisible module with finite corank, where as defined in* Section 8.1.6, $H^q_{ord}(X, ?)$ *is* $e_G H^q(X, ?)$ *if* $X = S_m$ *or* $T^N_{m,n}$ *and* $e^{\circ}_G H^q(X, ?)$ *if* $X = M_m$.

(4) *If* $\kappa \gg \mu_{m,n}$, $H^0_{ord}(S, \underline{\omega}^{\kappa} \otimes \mathbb{Q}_p/\mathbb{Z}_p)$ *(resp.* $H^0_{ord}(S_{\infty}, \underline{\omega}^{\kappa})$*) is isomorphic to* $H^0_{ord}(M, \underline{\omega}^{\kappa} \otimes \mathbb{Q}_p/\mathbb{Z}_p)$ *(resp.* $H^0_{ord}(M, \underline{\omega}^{\kappa})$*), where* $\kappa \gg \mu_{m,n}$ *means* $\kappa - \mu_{m,n}$ *is sufficiently interior of the positive Weyl chamber* $X_+(T)$;

(5) *Let* $\mathcal{V}^*_{ord}$ *be the Pontryagin dual* $\mathbb{Z}_p[[T(\mathbb{Z}_p)]]$*-module of* $e\mathcal{V}$ *(which is isomorphic to the* $\mathbb{Z}_p$*-dual of* eV*). Then* $\mathcal{V}^*_{ord}$ *is* $\mathbb{Z}_p[[\Gamma_T]]$*-free of finite rank, where* Γ_T *is the maximal* p*-profinite subgroup of* $T(\mathbb{Z}_p)$;

(6) *If* $\kappa \in X_+(T)$, $\mathcal{V}^*_{ord} \otimes_{\mathbb{Z}_p[[T(\mathbb{Z}_p)]],\kappa} \mathbb{Z}_p \cong \mathrm{Hom}_{\mathbb{Z}_p}(H^0_{ord}(S_{\infty}, \underline{\omega}^{\kappa}), \mathbb{Z}_p)$.

Proof. We need to check the conditions described in Section 8.1.7 to obtain the result as in the theorem. The boundedness condition (F) in Section 8.1.7 is verified in [H95], Theorem 1.1 for our unitary group U in the same manner as in the proof of Theorem 4.9 in the Hilbert modular case (using the generalized Eichler–Shimura map in Section 5.3), because $U(\mathbb{Z}_p) \cong GL_{m+n}(\mathbb{Z}_p)$. The hypothesis (Hp1) is clear because $S = Sh_K^{(p)}[\frac{1}{E}]$ is an affine scheme for our projective scheme $Sh_K^{(p)}$. The value $f|U_j(\mathfrak{p})(A, \lambda, i, \overline{\eta}^{(p)}, fil, \omega)$ is the sum (more precisely, the trace) of the value $f(A/C, \lambda', i', \overline{\eta'}^{(p)}, fil', \omega')$ for étale subgroups C of $A[p]$. Since the Hasse invariant is insensitive to étale isogeny (by its definition), the commutativity condition (C) in Section 8.1.7 holds. □

Although we restricted ourselves to the projective case here, a similar result can be obtained in more general settings, once we replace $\underline{\omega}^{\kappa}$ by its cuspidal subsheaf $\underline{\omega}_{\kappa}$. Even in the Hilbert modular case, as seen in Section 4.1.7, if the

base field is not equal to $\mathbb{Q}$, we can easily find an example that the cokernel of the inclusion $eH^0(S_{/W}, \underline{\omega}^\kappa_{/W}) \otimes_W \mathbb{F} \hookrightarrow eH^0(S_{/\mathbb{F}}, \underline{\omega}^\kappa_{/\mathbb{F}})$ is a nonzero finite module for κ densely populated in $X(T)$. Thus to obtain the result similar to the above theorem in the nonprojective case, we need to add some assumption, such as replacing $\underline{\omega}^\kappa$ by the cuspidal sheaf $\underline{\omega}_\kappa$ (although for $U(n,n)_{/\mathbb{Q}}$ and $GSp(2n)_{/\mathbb{Q}}$ it seems the result is true even for $\underline{\omega}^\kappa$). Here we state a precise result without proof only for the quasi-split unitary group $G = GU(n,n)$ for an inert prime p of F (when p splits in F, the result is basically included in Theorem 8.13 in the sense that we need to replace $\underline{\omega}^\kappa$ by the cuspidal sheaf $\underline{\omega}_\kappa$ and V and $\mathcal{V}$ by their cuspidal counterparts V_{cusp} and $\mathcal{V}_{cusp}$). Thus $D = F$, $V = F^{2n} = F^n \oplus F^n$, and $\langle (x_1, x_2), (y_1, y_2) \rangle = \mathrm{Tr}_{F/\mathbb{Q}}(x_1 {}^t\overline{y}_2 - x_2 {}^t\overline{y}_1)$. The lattice L can be given so that $L_p = O_p^n \oplus O_p^n$. The balanced maximal parabolic subgroup $P = \{\left(\begin{smallmatrix} * & * \\ 0_n & * \end{smallmatrix}\right)\}$ of the unitary group $U = U(n,n)$ has the reductive part isomorphic to $\mathrm{Res}_{F/\mathbb{Q}}(GL(n))$ via $\mathrm{Res}_{F/\mathbb{Q}}(GL(n)) \ni a \mapsto \mathrm{diag}[a, {}^t\overline{a}^{-1}] \in P$. We consider the diagonal torus $T \subset \mathrm{Res}_{F/\mathbb{Q}}(GL(n))$ and the unipotent radical N of the upper triangular Borel subgroup of $\mathrm{Res}_{F/\mathbb{Q}}(GL(n))$. Write N_G for the unipotent radical of the upper triangular Borel subgroup of G; so, $N = N_G/N_{n,n}$ for the unipotent radical $N_{n,n}$ of P. In the same way as in the Hilbert modular case (see Section 4.1.7), we define the space of cuspidal p-adic modular form $V_{cusp} := V^N_{cusp}$ (resp. its p-divisible counterpart: $\mathcal{V}_{cusp} := \mathcal{V}^N_{cusp}$) for the Igusa tower over $Sh^{(p)}_K$ for an open compact subgroup K of $G(\mathbb{A}^{(\infty)})$ maximal at p. In this case, the single $GL_n(O_p)$ acts on the tower (in place of $GL_n(O_\mathfrak{p}) \times GL_n(O_{\overline{\mathfrak{p}}})$ in the case where p splits).

Theorem 8.14 *Let $G_{/\mathbb{Q}}$ be the quasi-split unitary similitude group of signature (n,n). Suppose that p is inert in the imaginary quadratic field F. Let the notation be as in* Theorem 8.13. *Let $M_{/O_p}$ be the toroidal compactification of $Sh^{(p)}_K(GU, X)_{/O_p}$, and put $S = M[\frac{1}{E}]$ for a lift E of the Hasse invariant.*

(1) *We have an inclusion $\beta : \bigoplus_{\kappa \in X_+(T)} H^0(M_{/O_p}, \underline{\omega}^\kappa \otimes_{\mathbb{Z}_p} R) \hookrightarrow V \otimes_{\mathbb{Z}_p} R$ for $R = \mathbb{Z}_p$ and $\mathbb{Q}_p/\mathbb{Z}_p$ with the identity of q-expansion: $\beta(f)(q) = \ell_{can}(f)(q)$.*
(2) *$\mathrm{Im}(\beta_{cusp})[\frac{1}{p}] \cap V$ is dense in V_{cusp}, where β_{cusp} is the restriction of β to $\bigoplus_{\kappa \in X_+(T)} H^0(M_{/O_p}, \underline{\omega}_\kappa)$.*
(3) *Write $U(p) = \prod_{1 \le j \le n} u_j(p)$ for the standard Hecke operators at p of level p^∞, where, for $\alpha_j = \mathrm{diag}[1_j, p1_{n-j}, p^2 1_j, p1_{n-j}] \in G(\mathbb{Q}_p)$,*

$$u_j(p) = \mu_{n,n}(\alpha_j)^{-1}[N_G(\mathbb{Z}_p)\alpha_j N_G(\mathbb{Z}_p)].$$

Then $u_j(p)$ is well-defined over O_p if $\kappa \ge \mu_{n,n}$. Define the ordinary projector $e_G = \lim_{n\to\infty} U(p)^{n!}$ on the cuspidal subspace V_{cusp} of V. Then $H^0_{ord}(S, \underline{\omega}_\kappa \otimes_{\mathbb{Z}_p} \mathbb{Q}_p/\mathbb{Z}_p)$ is a p-divisible module with finite corank.
(4) *If κ is sufficiently regular, $H^0_{ord}(S, \underline{\omega}_\kappa \otimes_{\mathbb{Z}_p} \mathbb{Q}_p/\mathbb{Z}_p)$ is isomorphic to a direct summand $H^0_{ord}(M, \underline{\omega}_\kappa \otimes_{\mathbb{Z}_p} \mathbb{Q}_p/\mathbb{Z}_p)$ of $H^0(M, \underline{\omega}_\kappa \otimes_{\mathbb{Z}_p} \mathbb{Q}_p/\mathbb{Z}_p)$.*
(5) *Let $\mathcal{V}^*_{ord}$ be the Pontryagin dual $O_p[[T(\mathbb{Z}_p)]]$-module of $e_G(\mathcal{V}_{cusp})$. Then $\mathcal{V}^*_{ord}$ is $O_p[[\Gamma_T]]$-free of finite rank, where Γ_T is the maximal p-profinite subgroup of $T(\mathbb{Z}_p)$.*

(6) *The* $O_p[[T(\mathbb{Z}_p)]]$*-module* $\mathcal{V}^*_{ord}(\cong \mathrm{Hom}_{O_p}(e_G(\mathcal{V}_{cusp}), O_p))$ *satisfies*

$$\mathcal{V}^*_{ord} \otimes_{O_p[[T(\mathbb{Z}_p)]],\kappa} O_p \cong \mathrm{Hom}_{O_p}(H^0_{ord}(S_\infty, \underline{\omega}_\kappa), O_p)$$

if $\kappa \geq 0$, *where* $H^0_{ord}(S_\infty, \underline{\omega}_\kappa) = \varprojlim_m H^0_{ord}(S_m, \underline{\omega}_\kappa)$. *Similarly, if* $\kappa \gg \mu_{n,n}$, *we have*

$$\mathcal{V}^*_{ord} \otimes_{O_p[[T(\mathbb{Z}_p)]],\kappa} O_p \cong \mathrm{Hom}_{O_p}(H^0_{ord}(M, \underline{\omega}_\kappa), O_p).$$

The proof for the cuspidal subsheaves in the nonprojective cases is basically the same as in the projective case except for an extra effort necessary to verify (Hp1), because in the nonprojective case, the ordinary locus S_m of the smooth toroidal compactification is no longer affine. This can be done via q-expansion at all cusps of the Igusa tower (in the same manner as in the Hilbert modular case, Section 4.1.7). As we show in the following section, the Igusa tower T°_∞ over a geometrically irreducible component S°_1 of S_1 for general unitary groups may not be geometrically irreducible, although each irreducible component of T°_∞ contains a cusp. So, the naive use of q-expansion only at one cusp (often ∞) in the elliptic modular case (done in Section 4.2.9) to show the integrality of the Hecke operator no longer works in general, but we need to use one cusp at each component. In any case, as we have done in the split case, one may use the Serre–Tate deformation coordinates in place of the q-expansion in order to prove the integrality of $U(\mathfrak{p})$ in the general case. See [H02] for details of the arguments in the nonprojective case.

8.4 Irreducibility of Igusa Towers

We sketch a proof of the irreducibility of the generalized Igusa tower by using the determination of the automorphism group of the arithmetic automorphic function field by Shimura and others (cf. Theorem 7.7). In this section, *we deal with a general CM field* F *not necessarily imaginary quadratic* and study general groups GU of types A and C admitting Shimura varieties of PEL type.

The method we use is classical (based on the determination of the decomposition group of p-adic valuations of the arithmetic automorphic function field) and goes back to a work of Deuring [Du] for the elliptic modular function field. By this result, the q-expansion principle holds for p-adic modular forms on these groups. Shimura has now completed his computation of Fourier coefficients of Eisenstein series of quasi-split groups having a balanced parabolic subgroup of the form $\{\left(\begin{smallmatrix} * & * \\ 0 & * \end{smallmatrix}\right)\}$ ([EPE] 18.14 and [AAF] 16.9-11). We should therefore be able to construct the p-adic Eisenstein measure for such quasi-split unitary groups, although we do not cover this topic in this book. Since we can identify the p-adic completion of the space of classical automorphic forms with the space of p-adic automorphic forms by q-expansion, the measure (if constructed) is now verified to have values in the space of p-adic automorphic forms and could be specialized at a point of the Shimura variety

to obtain certain p-adic automorphic L-functions (this point is not verified in the construction by Panchishkin in [P] of a "naive" Eisenstein measure for Siegel modular varieties).

In this section, $\mathcal{W}$ is the pullback by $i_p : \overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_p$ of the p-adic integer ring of the maximal unramified extension of $\mathbb{Q}_p$ in $\overline{\mathbb{Q}}_p$; so, it is a strict Henselization of $\mathbb{Z}_{(p)}$ inside $\overline{\mathbb{Q}}$, and its residue field $\mathbb{F}$ is an algebraic closure of $\mathbb{F}_p$. The valuation of $\mathcal{W}$ and $W = \varprojlim_n \mathcal{W}/p^n\mathcal{W}$ is written as ord_p. For the reflex field E of (GU, X), we write $\mathcal{V}$ for $\mathcal{W} \cap E$ and κ for the residue field of $\mathcal{V}$.

8.4.1 Irreducibility and p-Decomposition Groups

Let the notation be as in Section Section 7.1, in particular 7.1.3, and we suppose the simply connectedness condition (SC) for G_1; so, we exclude groups of type D. Probably Case D could be treated similarly, but due to the lack of the strong approximation theorem for $G_1(\mathbb{A}^{(\infty)})$, the description of the connected components should be more complicated.

By the assumption (SC), the strong approximation theorem holds for $G_1(\mathbb{A}^{(\infty)})$, and the geometrically connected components of $Sh(GU, X)_{/E}$ over E are indexed by $\pi_0\pi(Z^{GU})$ for the cocenter $Z^{GU} := GU/G_1$ of GU (see Section 7.2.3 for $\pi_0\pi(Z^{GU})$). Note that the group GU satisfies (D1–4), (SC), and (CT) if the involution ι of $C = \mathrm{End}_D(V)$ is not definite (positive or negative). Hereafter we assume that the involution ι is indefinite; so, $\dim Sh(GU, X) > 0$. Under this assumption, Theorem 7.7 holds.

Let $\mathfrak{K}$ be the function field of a geometrically irreducible component of $Sh(G, X)_{/E}$ for the reflex field E of (G, X). The field $\mathfrak{K}$ is called the *arithmetic automorphic function field* for G. We show that all elements in the decomposition group of a well-chosen p-adic valuation of the field $\mathfrak{K}$ have their p-components in a maximal parabolic subgroup P of GU (in the Siegel modular case, P is given by the Siegel parabolic subgroup; see Theorem 6.27). For the reductive quotient M of P, $M(\mathbb{Z}_p)$ gives a subgroup of the automorphism group of the Igusa tower over the Shimura variety. If the projection $\overline{\nu} : GU(\mathbb{A}^{(\infty)}) \to \pi_0\pi(Z^{GU})$ is nontrivial over $M(\mathbb{Z}_p)$, we cannot expect that the Igusa tower over an irreducible base is irreducible and has Galois group $M(\mathbb{Z}_p)$, because some of the components belong to different generic geometrically irreducible components. We expect therefore that the Galois group of the irreducible component of the tower is isomorphic to $M_1(\mathbb{Z}_p)$ for $M_1 = M \cap G_1$.

8.4.2 Closed Immersion into the Siegel Modular Variety

Hereafter we use the symbols (G, X_G) and (GU, X_{GU}) exclusively for those defined in Section 7.1.3 (particularly, (7.1)). Write $(GSp(2g)_{/\mathbb{Q}}, \mathcal{X} \cong \mathfrak{H}_g \sqcup \overline{\mathfrak{H}}_g)$ for the pair described in (C) in Example 7.6. For simplicity, we write

$$Sh_G = Sh(G, X_G),\ Sh_{GU} = Sh(GU, X_{GU}),\ Sh_{GSp(2g)} = Sh(GSp(2g), \mathcal{X}).$$

Note that $\mathcal{G}(GSp(2g),\mathcal{X}) = GSp_{2g}(\mathbb{A})_+$ by Theorem 6.26. We write $\mathfrak{F}$ for the arithmetic automorphic function field $\mathbb{Q}^{cyc}(\mathfrak{M})$ in Theorem 6.26; so, $GSp_{2g}(\mathbb{A}^{(\infty)})/Z(\mathbb{Q}) \cong \mathrm{Aut}(\mathfrak{F})$. Similarly we write $\mathfrak{F}^{(p)} = \mathbb{Q}^{(p)}(\mathfrak{M}^{(p)})$ for the subfield of $\mathfrak{F}$ giving the function field of the generic fiber of the p-integral model of level prime-to-p. Then $\mathrm{Aut}(\mathfrak{F}^{(p)}) = GSp_{2g}(\mathbb{A}^{(p\infty)})/Z(\mathbb{Z}_{(p)})$.

We write $\mathfrak{K}$ (resp. $\mathfrak{K}^{(p)}$) for the arithmetic function field of the neutral component of Sh_{GU} (resp. $Sh^{(p)}_{GU}$). By our assumptions (SC) and (CT), we have the group $\mathcal{G}(GU, X_{GU})$ for the pair (GU, X_{GU}) (as in Theorem 7.7) acting on the field $\mathfrak{K}$ with constant field $\mathfrak{k} = \mathfrak{K} \cap \overline{\mathbb{Q}}$ which is an abelian extension of the reflex field E. In particular, $\tau(\mathcal{G}(GU, X_{GU})) \cong \mathcal{G}(GU, X_{GU})/GU(\mathbb{R})_+\overline{Z(\mathbb{Q})}$ is a subgroup of finite index of $\mathrm{Aut}(\mathfrak{K})$.

Since the Shimura variety Sh_G of PEL type carries the universal abelian scheme $\underline{A}_G = (\mathcal{A}_G, \boldsymbol{\lambda}_G, \boldsymbol{\eta}_G)$, by the universality of the triple $\underline{A}_{GSp(2g)/\mathbb{Q}}$ over $Sh_{GSp(2g)}$, we have a unique morphism

$$i : Sh_{G/E} \to Sh_{GSp(2g)/E} \quad \text{for } g = \dim_{Sh_{GU}} \mathcal{A}_{GU}$$

such that $i^*\underline{A}_{GSp(2g)} \approx \underline{A}_G$. Our construction of Sh_G in Sections 7.1.2 and 7.1.3 tells us the finiteness and unramifiedness of this morphism (Corollary 7.2).

Since $h_x \in X_G$ gives rise to the complex structure on $V_{\mathbb{R}} = V \otimes_{\mathbb{Q}} \mathbb{R}$ and the fiber A of $\mathcal{A}_G$ at x has expression $A(\mathbb{C}) = V_{\mathbb{R}}/L$ as a complex torus, the morphism i is described as $h_x \mapsto \phi \circ h_x$ at the level of the symmetric space for an embedding $\phi : G_{/\mathbb{R}} \hookrightarrow GSp(2g)_{/\mathbb{R}}$, where G is the subgroup of GU with rational similitude (defined in Section 7.1.1). Recall that $X^+_G = X^+_{GU}$ for which we write X^+. The natural morphism $i_G : Sh_G \to Sh_{GU}$ with $i_G^*\underline{A}_{GU} \approx \underline{A}_G$ identifies the neutral component of Sh_G and that of Sh_{GU}. This also follows from the fact $\tau(\mathcal{G}(GU, X_{GU})) = \overline{\mathcal{E}}(G, X_G)$ (for $\overline{\mathcal{E}}(G, X_G)$ in [MiS] 4.2) inside

$$\frac{G(\mathbb{A}^{(\infty)})}{\overline{Z_G(\mathbb{Q})}} *_{G(\mathbb{Q})_+} G^{ad}(\mathbb{Q})^+ \subset \frac{GU(\mathbb{A}^{(\infty)})}{\overline{Z_{GU}(\mathbb{Q})}} *_{GU(\mathbb{Q})_+} GU^{ad}(\mathbb{Q})^+ = \frac{GU(\mathbb{A}^{(\infty)})}{\overline{Z_{GU}(\mathbb{Q})}},$$

because the left-hand side of the above inclusion is the closure of the image of $G(\mathbb{A}^{(\infty)})$ in $\frac{GU(\mathbb{A}^{(\infty)})}{\overline{Z_{GU}(\mathbb{Q})}}$. Thus $\mathcal{G}(GU, X_{GU})$ acts on the neutral component of Sh_G. The embedding ϕ is obtained by regarding $\alpha \in G(\mathbb{Q})$ as a $\mathbb{Q}$-linear automorphism $\phi(\alpha)$ of V preserving the symplectic form $\langle \cdot, \cdot \rangle$ in (L2) of Section 7.1.1; so, ϕ takes $G(\mathbb{Q})$ into $GSp(2g)$. Thus ϕ is defined over $\mathbb{Q}$. In other words, ϕ is induced by $C = \mathrm{End}_D(V) \hookrightarrow \mathrm{End}_{\mathbb{Q}}(V)$, and $\phi(g)(i(x)) = i(g(x))$ for $x \in X^+$ and $g \in GU(\mathbb{R})_+$.

Using the description of complex points $Sh_G(\mathbb{C})$ in (7.7), it is easy to verify that i is an embedding $Sh_G(\mathbb{C}) \hookrightarrow Sh_{GSp(2g)}(\mathbb{C})$ of complex manifolds. Thus $i : Sh_G \to Sh_{GSp(2g)}$ is a closed immersion ([D1] 5.4) defined over $E = E(G, X_G) = E(GU, X_{GU})$ compatible with reciprocity laws in Theorem 7.7 (4). In particular the reflex field E and the constant field $\mathbb{Q}^{cyc}$ of the arithmetic function field $\mathfrak{F}$ for $(GSp(2g), \mathcal{X})$ are linearly disjoint over $k = \mathbb{Q}^{cyc} \cap E$; so, we

may identify $\tau(GSp_{2g}(\mathbb{A})_+)$ with an open subgroup of $\mathrm{Aut}(E\mathfrak{F})$, where $E\mathfrak{F} = E \otimes_k \mathfrak{F}$ (the arithmetic automorphic function field for the neutral component of $Sh_{GSp(2g)/E}$).

We now choose a lattice L satisfying the conditions (L1–2) in Section 7.1. We choose a $\mathbb{Z}$-base of L so that the matrix of the alternating form $\langle \cdot, \cdot \rangle$ is given by J_g. Then we may assume that

$$\phi^{-1}(GSp_{2g}(\mathbb{Z}_p)) = G(\mathbb{Z}_p) = \{x \in G(\mathbb{Q}_p) | x_p L_p = L_p\}.$$

Recall that the p-integral model $Sh_G^{(p)}$ is defined over $\mathcal{V} = \mathcal{W} \cap E$. The morphism i induces a finite unramified morphism of the p-integral models

$$i : Sh^{(p)}_{G/\mathcal{V}} \to Sh^{(p)}_{GSp(2g)/\mathcal{V}}.$$

Since ϕ brings the connected component $G(\mathbb{R})^+$ into $GSp_{2g}(\mathbb{R})^+$, i brings the geometrically irreducible component (the neutral component) $\mathfrak{M}_G^{(p)} = \mathfrak{M}_{GU}^{(p)}$ of $Sh_G^{(p)}$ containing the image of $X^+ \times 1$ into the neutral component $\mathfrak{M}^{(p)} = \varprojlim_{p \nmid M} \mathfrak{M}_M^J$ of $Sh^{(p)}_{GSp(2g)/\mathbb{Q}}$.

Supposing the condition (ord) requiring that the neutral component $\mathfrak{M}_G^{(p)} \otimes_{\mathcal{V}} \mathbb{F}$ has dense p-ordinary locus S_1, we look at the stalk $\mathcal{O}_{S_1,x}$ at a closed point $x \in S_1(\overline{\mathbb{F}}_p)$ in the neutral component of S_1. The formal completion $\widehat{S}_{1,x} = Spf(\widehat{\mathcal{O}}_{S_1,x})$ along x is the formal deformation space of an ordinary abelian variety A_x sitting on x. By the description of the deformation space in Sections 8.2.4 and 8.2.5 (see the following subsection for more details of this point), we conclude that i induces an embedding of $\widehat{S}_{1,x}$ into the formal completion along $i(x)$ of the Siegel modular variety. Since i is unramified over $Sh^{(p)}_{GSp(2g)}$ (see Corollary 7.2) and i is generically a closed immersion, i above is a closed immersion over $\mathcal{V}$ of $\mathfrak{M}_G^{(p)}$ into the ordinary locus of $\mathfrak{M}^{(p)}_{/\mathbb{F}}$ for the residue field $\mathbb{F} = \overline{\mathbb{F}}_p$ of $\mathcal{W}$, although we only need the fact that i is a closed immersion on an open dense subscheme of $Sh_G^{(p)}$ faithfully flat over $\mathcal{V}$. In any case, the function field of the geometrically irreducible component of the open dense subscheme coincides with the one of the entire neutral component, over E or over κ, where we write κ for the residue field of $\mathcal{V}$.

Pick $x \in Sp_{2g}(\mathbb{A}^{(\infty)})$ with $x_p \in GSp_{2g}(\mathbb{Z}_p)$. Then the scheme automorphism $\tau(x) \in \mathrm{Aut}(Sh_{GSp(2g)})$ induces an automorphism of the p-integral model $Sh^{(p)}_{GSp(2g)}$, which is in turn induced by an isogeny

$$\alpha(x) : \mathcal{A}_{GSp(2g)} \to \tau(x)^* \mathcal{A}_{GSp(2g)} \text{ such that } \alpha(x) \circ \boldsymbol{\eta}^{(p)} = \boldsymbol{\eta}^{(p)} \circ x$$

by the universality of $\mathcal{A}_{GSp(2g)}$. Since an isogeny preserves p-ordinarity, $\tau(x)$ leaves stable the p-ordinary locus $\mathfrak{S}_1$ of the neutral component $\mathfrak{M}^{(p)}_{/\mathbb{F}} \subset Sh^{(p)}_{GSp(2g)} \otimes_{\mathbb{Z}_{(p)}} \mathbb{F}$. As we have seen in the proof of Theorem 6.28, if $\tau(x)$

is in the decomposition group $\mathcal{D}_p$ of the valuation v on $\mathfrak{F}$ defined there, $\tau(x)$ preserves (over the p-ordinary locus $\mathfrak{S}_1$) the connected component $\mathcal{A}_{GSp(2g)}[p^\infty]^\circ \subset \mathcal{A}_{GSp(2g)}[p^\infty]$. This property is equivalent to the fact that x_p is in the Siegel parabolic subgroup $P_g(\mathbb{Z}_p)$ for $P_g = \{\left(\begin{smallmatrix} * & * \\ 0 & * \end{smallmatrix}\right)\}$. Since $\alpha(x)^{-1}(\mathcal{A}_{GSp(2g)}[p^r]^\circ)$ is a locally free subgroup scheme of rank $(p^r)^g$, if $\alpha(x)$ sends the fiber of the connected component $\mathcal{A}_{GSp(2g)}[p^\infty]^\circ$ (of $\mathcal{A}_{GSp(2g)}[p^\infty]$) at one point in $\mathfrak{S}_1$ into another fiber of $\mathcal{A}_{GSp(2g)}[p^\infty]^\circ$, it preserves the entire $\mathcal{A}_{GSp(2g)}[p^\infty]^\circ$, because $\alpha(x)$ is induced by a global isogeny. This fact is important in the determination of the p-decomposition group.

Since the Igusa tower can be created as a subscheme of a self product of copies of the étale quotient of $\mathcal{A}^{ord}_{GU}[p^\alpha]$ for the ordinary locus $\mathcal{A}^{ord}_{GU}$ of $\mathcal{A}_{GU}$ (see (8.1)), the embedding i extends to the embedding of the Igusa tower of G into that of $GSp(2g)$, and $\phi^{-1}(P_g)$ is a maximal parabolic subgroup of G well-defined over $\mathbb{Z}_p$ which gives rise to a unique maximal parabolic subgroup $P \subset GU$ defined over $\mathbb{Z}_p$ (i.e., the normalizer of $\phi^{-1}(P_g)$ in GU).

By Theorem 7.5, we have $Sh^{(p)}_{GU} \otimes_{\mathcal{V}} E = Sh_{GU}/GU(\mathbb{Z}_p)$. From this fact, we have a unique geometrically irreducible component $\mathfrak{M}_G = \mathfrak{M}_{GU}$ (inside $i_G(Sh_G)$) of Sh_{GU} covering $\mathfrak{M}^{(p)}_G \times_{\mathcal{W}} \mathcal{K}$ for the quotient field $\mathcal{K}$ of $\mathcal{W}$, which is the neutral component of Sh_{GU}. Then i brings $\mathfrak{M}_{GU}$ into the neutral component $\mathfrak{M} = \varprojlim_M \mathfrak{M}^J_M \subset Sh_{GSp(2g)}$ over $\overline{\mathbb{Q}}$.

8.4.3 Description of a p-Decomposition Group

Let $\mathfrak{V}_v$ be the valuation ring of $\mathfrak{F}$ with maximal ideal $\mathfrak{m}_v$ with respect to the valuation v introduced in the proof of Theorem 6.27. Thus for any $f \in \mathfrak{F}$, $p^\alpha f \in \mathfrak{V}_v$ for a suitable $\alpha \in \mathbb{Z}$. For Siegel modular forms f defined over $\mathbb{Q}^{cyc}$, we have $v(f) = \mathrm{Inf}_\xi\, ord_p(a(\xi, f))$ for the q-expansion coefficients $a(\xi, f)$ at the cusp ∞. We then consider the subring $\mathfrak{V}_{GU}$ of $\mathfrak{F}$ made of meromorphic functions that are well-defined over $i(\mathfrak{M}_{GU})$. Thus $\mathfrak{V}_{GU}$ is the localization of the sheaf of ideals giving the closed subscheme $i(\mathfrak{M}_{GU})$ inside $\mathfrak{M}$, and $\mathfrak{V}_{GU}$ is a local ring with maximal ideal $\mathfrak{Q} = \left\{f \in \mathfrak{V}_{GU} \middle| f|_{i(\mathfrak{M}_{GU})} = 0\right\}$. The elements $f \in \mathfrak{V}_{GU}$ are smooth over an open dense subscheme in the image of i.

Let $\mathfrak{V} = \mathfrak{V}_v \cap \mathfrak{V}_{GU}$ and $X^{(p)} = X \cap \mathfrak{F}^{(p)}$ for any subring X of $\mathfrak{F}$. Since $\mathfrak{V}_v$ and $\mathfrak{V}_{GU}$ are normal, respectively, over $\mathfrak{V}^{(p)}_v$ and $\mathfrak{V}^{(p)}_{GU}$, $\mathfrak{V}$ is normal over $\mathfrak{V}^{(p)}$. The scheme $\mathfrak{M}^{(p)}_{/\mathcal{W}}$ has a unique irreducible special fiber $\mathfrak{M}^{(p)}_{/\mathbb{F}} = \mathfrak{M}^{(p)} \otimes_{\mathcal{W}} \mathbb{F}$. As we show in the following subsection via the Serre–Tate local coordinates, at each ordinary CM point $x \in \mathfrak{M}^{(p)}_{GU}(\mathbb{F})$, the vertical modulo p special fiber $\mathfrak{M}^{(p)}_{/\mathbb{F}}$ and the horizontal closed subscheme $\mathfrak{M}^{(p)}_{GU/\mathcal{W}}$ of $\mathfrak{M}^{(p)}_{/\mathcal{W}}$ intersect transversally (i.e., the intersection is a local complete intersection at all ordinary points). Thus $\mathfrak{m}_v \cap \mathfrak{V}^{(p)} = (p)$ and $\mathfrak{Q} \cap \mathfrak{V}^{(p)}$ are two distinct prime ideals in $\mathfrak{V}^{(p)}$, and hence we have two prime ideals $\mathfrak{p} = \mathfrak{m}_v \cap \mathfrak{V}$ and $\mathfrak{q} = \mathfrak{V} \cap \mathfrak{Q}$ of $\mathfrak{V}$. The localization of $\mathfrak{V}$ at $\mathfrak{p}$ (resp. $\mathfrak{q}$) is equal to $\mathfrak{V}_v$ (resp. $\mathfrak{V}_{GU}$). By definition, $\mathfrak{V}/\mathfrak{q}$ is a subring of $\mathfrak{K}$ with $\mathfrak{K} = \bigcup_n p^{-n}(\mathfrak{V}/\mathfrak{q})$. Let $\mathfrak{m}_{GU}$ be the prime ideal $(\mathfrak{p}+\mathfrak{q})/\mathfrak{q}$ in $\mathfrak{V}/\mathfrak{q} \subset \mathfrak{K}$. Let

$$\mathcal{D}_{GU,v} = \left\{\sigma \in \tau(\mathcal{G}(GU,X))\middle|\ \sigma(\mathfrak{m}_{GU}) = \mathfrak{m}_{GU}\right\}$$
$$\mathcal{D}_v = \left\{\sigma \in \tau(GSp_{2g}(\mathbb{A})_+) \cap \mathrm{Aut}(E\mathfrak{F}/E)\middle|\ \sigma(\mathfrak{m}_v) = \mathfrak{m}_v\right\}$$
$$\mathcal{D}_{\mathfrak{Q}} = \left\{\sigma \in \tau(GSp_{2g}(\mathbb{A})_+) \cap \mathrm{Aut}(E\mathfrak{F}/E)\middle|\ \sigma(\mathfrak{Q}) = \mathfrak{Q}\right\}.$$

Since the localization of $\mathfrak{V}$ at $\mathfrak{p}$ (resp. $\mathfrak{q}$) is equal to $\mathfrak{V}_v$ (resp. $\mathfrak{V}_{GU}$), we see that

$$\mathcal{D}_v \cap \mathcal{D}_{\mathfrak{Q}} = \left\{\sigma \in \tau(GSp_{2g}(\mathbb{A})_+) \cap \mathrm{Aut}(\mathfrak{V}/\mathcal{V})\middle|\ \sigma(\mathfrak{q}) = \mathfrak{q} \ \text{ and } \ \sigma(\mathfrak{p}) = \mathfrak{p}\right\}.$$

The subgroup $\mathcal{D}_v$ is an open subgroup of $\mathcal{D}_p$ (in Theorem 6.28) of finite index, and they are equal if $E \cap \mathbb{Q}^{cyc} = \mathbb{Q}$. We want to show first

Theorem 8.15 *Suppose that the group GU admits the Shimura variety of PEL-type $Sh^{(p)}_{GU}$ described in* (A) *and* (C) *of* Example 7.6 *and that the Hasse invariant does not vanish identically on $Sh^{(p)}_{GU} \otimes_{\mathcal{W}} \mathbb{F}$. Then τ brings*

$$\left\{g \in \mathcal{G}(GU,X)\middle|\phi(g_p) \in P_g(\mathbb{Q}_p) \text{ and } \sigma(g) \text{ preserves the valuation } ord_p \text{ on } \mathfrak{k}\right\}$$

onto $\mathcal{D}_{GU,v}$, where $\sigma : \mathrm{Aut}(\mathfrak{K}) \to \mathrm{Gal}(\mathfrak{k}/E)$ is the restriction map. In particular, $\tau^{-1}(\mathcal{D}_{GU,v}) = \phi^{-1}(\mathcal{D}_v)$, and $\mathcal{D}_{GU,v}$ is isomorphic to $\tau(\phi(\mathcal{G}(GU,X))) \cap \mathcal{D}_v$ in $\mathrm{Aut}(E\mathfrak{F}/E)$.

Proof. We prove the first assertion (because the last one follows from the first). The function field $\mathfrak{K} = \mathfrak{V}_{GU}/\mathfrak{Q}$ is characterized by the reciprocity law (4) of Theorem 7.7. Indeed, $\tau(g)$ on $\mathfrak{K}$ is induced by $\eta \mapsto \eta \circ g$ for the level structure η of test objects. This definition of τ for the Siegel modular function field is the same for Siegel modular test objects, which implies that

$$({}^{\tau(\phi(g))}f) \circ i = {}^{\tau(g)}(f \circ i) \ \text{ for each } f \in \mathfrak{V}_{GU} \text{ and } g \in \overline{\mathcal{G}}(GU,X). \tag{8.28}$$

For each special point $x \in X^+$ and the reciprocity map $r_x : \mathrm{Gal}(E_x^{ab}/E_x) \to T_x(\mathbb{A})/\overline{T}_x(\mathbb{Q})$, $i(x) \in \mathcal{X}^+$ is a special point for $\mathfrak{M}$, and by definition in Section 7.2.2, we find $r_{i(x)} = \phi \circ r_x$. Then we confirm from the local reciprocity law (rc1–2) that

$$({}^{\tau(\phi(r_x(\sigma)))}f) \circ i(x) = {}^{\tau(r_x(\sigma))}(f \circ i)(x)$$

for each $f \in \mathfrak{V}_{GU}$ and $\sigma \in \mathrm{Gal}(E_x^{ab}/E_x)$. Since $T_x(\mathbb{A})$ generates $GU(\mathbb{A})$ (moving around all special points $x \in X^+$), from the density of special points in $\mathfrak{M}_{GU}$, we confirm (8.28). Therefore i induces a well-defined homomorphism $\phi : \tau(\mathcal{G}(GU,X)) \to \mathrm{Aut}(\mathfrak{F}E/E)$.

From (8.28), we find $\mathcal{D}_{\mathfrak{Q}} \supset \tau(\phi(\mathcal{G}(GU,X)))$. Therefore, $\tau(\phi(g))$ preserves the prime ideal $\mathfrak{Q} \subset \mathfrak{V}_{GU}$, and for each $\sigma \in \mathcal{D}_v \cap \mathcal{D}_{\mathfrak{Q}}$, we have $\sigma(\mathfrak{p} + \mathfrak{q}) \subset \mathfrak{p} + \mathfrak{q}$, which shows that σ as an automorphism of $\mathfrak{K}$ preserves $\mathfrak{m}_{GU}$. We get

$$\phi(\mathcal{D}_{GU,v}) \supset \left(\mathcal{D}_v \cap \tau(\phi(\mathcal{G}(GU,X)))\right).$$

Recall the lattice L of V (we have chosen as above) satisfying the conditions (L1–2) in Section 7.1 for GU and $GSp(2g)$ at the same time. We have the maximal compact subgroup $GU(\mathbb{Z}_p) = \{x \in GU(\mathbb{Q}_p) | xL_p = L_p\}$ of $GU(\mathbb{Q}_p)$ and $G(\mathbb{Z}_p) = GU(\mathbb{Z}_p) \cap G(\mathbb{Q}_p)$ of $G(\mathbb{Q}_p)$, and we may assume that $G(\mathbb{Z}_p) = \phi^{-1}(GSp_{2g}(\mathbb{Z}_p))$ (choosing a suitable $\mathbb{Z}$-base of L). If a prime factor $\mathfrak{p}_0$ of p in $O_0 = O_{F_0}$ splits into $\mathfrak{p}\overline{\mathfrak{p}}$ in O, we have the $\mathfrak{p}_0$ factor of $GU(\mathbb{Z}_p)$ isomorphic to $GL_n(O_\mathfrak{p}) \times O_\mathfrak{p}^\times$ for a suitable n. If $\mathfrak{p}_0 | p$ is a prime factor of p in O_0 giving a prime $\mathfrak{p} = \mathfrak{p}_0 O$ of O, the $\mathfrak{p}_0$-factor of $GU(\mathbb{Z}_p)$ is isomorphic to $GU_0(O_{0,\mathfrak{p}})$ (see Lemma 8.10), where GU_0 is the quasi-split unitary group defined over $O_0 \otimes_\mathbb{Z} \mathbb{Z}_{(p)}$ given by

$$GU_0(R) = \left\{\alpha \in GL_{2m(\mathfrak{p})}(R \otimes_{O_0} O) \middle| \alpha J_{m(\mathfrak{p})} {}^t\alpha^c = \nu(\alpha) J_{m(\mathfrak{p})} \quad \text{for } \nu(\alpha) \in R^\times\right\} \tag{8.29}$$

for $O_0 \otimes_\mathbb{Z} \mathbb{Z}_{(p)}$-algebras R ($m(\mathfrak{p})$ is independent of the inert $\mathfrak{p}$). Similarly we define

$$U_0(R) = \left\{\alpha \in GU_0(R) \middle| \alpha J_n {}^t\alpha^c = J_n\right\}. \tag{8.30}$$

In other words, $GU(\mathbb{Z}_p)$ is a hyperspecial maximal compact subgroup of $GU(\mathbb{Q}_p)$ in the sense of [Tt1].

Pick $\sigma \in \mathcal{D}_{GU,v}$. Suppose that $\sigma = \tau(g)$ with $g_p \in GU(\mathbb{Z}_p)$. Thus σ preserves the function field $\mathfrak{K}^{(p)} = \mathfrak{K}^{GU(\mathbb{Z}_p)}$ of the generic fiber of the Kottwitz model $Sh^{(p)}_{GU}$. Since $Sh^{(p)}_K(GU, X_{GU})_{/\mathcal{W}\cap\mathfrak{k}}$ carries a universal abelian scheme $\mathcal{A}_{GU}$ having the dense open ordinary locus $\mathcal{A}^{ord}_{GU}$, we have the isogeny $\mathcal{A}^{ord}_{GU} \approx i^*\mathcal{A}^{ord}_{GSp(2g)}$. In particular, $\sigma \in \mathcal{D}_{GU,v}$ preserves the p-ordinary locus. The exact sequence $\mathcal{A}^{ord}_{GU}[p^n]^\circ \hookrightarrow \mathcal{A}^{ord}_{GU}[p^n] \twoheadrightarrow \mathcal{A}^{ord}_{GU}[p^n]^{et}$ is the pullback of the corresponding one $\mathcal{A}^{ord}_{GSp(2n)}[p^n]^\circ \hookrightarrow \mathcal{A}^{ord}_{GSp(2n)}[p^n] \twoheadrightarrow \mathcal{A}^{ord}_{GSp(2n)}[p^n]^{et}$ by i. As we have already remarked, if the action of $x \in GSp_{2g}(\mathbb{Z}_p \times \mathbb{A}^{(\infty)})$ sends one fiber of the above exact sequence to another fiber of the same exact sequence (over the given connected component of $Sh_{GSp(2g)}$), we know that $x_p \in P_g(\mathbb{Z}_p)$. This tells us that $\phi(g_p) \in P_g(\mathbb{Z}_p)$ for the Siegel parabolic subgroup $P_g = \{\left(\begin{smallmatrix} * & * \\ 0 & * \end{smallmatrix}\right)\}$. When $g_p \notin GU(\mathbb{Z}_p)$, by choosing an element $\pi \in P_g(\mathbb{Q})$, we find $\phi(g)\pi$ has a p-component in $GSp_{2g}(\mathbb{Z}_p)$, because of the density of $P_g(\mathbb{Q})$ in $P_g(\mathbb{Q}_p)$ and the Iwasawa decomposition $GSp_{2g}(\mathbb{Q}_p) = GSp_{2g}(\mathbb{Z}_p) \cdot P_g(\mathbb{Q}_p)$ (see [EPE] 7.2). Still $\tau(\phi(g)p)$ keeps some fibers of the étale-connected exact sequence; so, by the same argument, we find that $\phi(g_p) \in P_g(\mathbb{Q}_p)$. This combined with Theorem 6.28 proves the desired result. □

8.4.4 Irreducibility Theorem in Cases A and C

We describe the parabolic subgroup $P \subset GU(\mathbb{Q}_p)$ given by $\phi^{-1}(P_g)$. First suppose that we are in Case A. Recall the ordinarity condition we assumed in addition to (unr) in Section 7.1.1:

(ord) *The Hasse invariant of $\mathcal{A}_{GU}$ does not vanish identically on $\mathfrak{M}^{(p)}_{GU/\mathbb{F}}$.*

See Lemma 8.10 for an explicit description of GU satisfying this condition. Then the signature of the Hermitian pairing $\langle\ ,\ \rangle$ on $L \otimes_\sigma \mathbb{R}$ depends only on the p-adic place v_σ of F induced by $i_p \circ \sigma$. Let us explain this. Take the idempotent $\varepsilon = \mathrm{diag}[1, 0, \ldots, 0] = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0_{n-1} \end{smallmatrix}\right) \in O_{D,p}$ with $\varepsilon^* = \varepsilon$, and regard it as the endomorphism of the Barsotti–Tate group $\mathcal{A}_{GU}[p^\infty]$. For the prime ideal $\mathfrak{p}$ of O corresponding to v, if we write $m(\mathfrak{p})$ for the generic rank of the maximal étale quotient of $\varepsilon \cdot \mathcal{A}_{GU}[\mathfrak{p}^\infty]$, the signature is given by $(m(\mathfrak{p}), m(\mathfrak{p}^c))$ for the automorphism c of F induced by complex conjugation. Here $\mathfrak{p}^c$ is induced by $i_p \circ \sigma c$. We split the set $\mathcal{S}$ of all p-adic places of F into a disjoint union $\Sigma_{sp} \sqcup \Sigma_{sp}^c \sqcup \Sigma_i$ so that $\Sigma_{sp} \sqcup \Sigma_{sp}^c$ is made up of all prime factors of p split over F_0, Σ_i is made up of the rest (inert primes), and $m(\mathfrak{p}) \geq m(\mathfrak{p}^c)$ for all $\mathfrak{p} \in \Sigma_{sp}$. For $\mathfrak{p} \in \Sigma_i$, $m(\mathfrak{p}) = m(\mathfrak{p}^c)$ because $\mathfrak{p} = \mathfrak{p}^c$ (see Lemma 8.10). We write this signature as $(m(\mathfrak{p}), m(\mathfrak{p}^c))$ for each p-adic place $\mathfrak{p} \in \mathcal{S}$. Then $n = m(\mathfrak{p}) + m(\mathfrak{p}^c)$ is independent of $\mathfrak{p} \in \mathcal{S}$. We write $\mathfrak{p}_0$ for the restriction $\mathfrak{p}$ to the (totally real) subfield F_0 of F fixed by c. Let $\Sigma = \Sigma_{sp} \sqcup \Sigma_i$. Then Σ is in bijection with the set $\mathcal{S}_0$ of all p-adic places of F_0 (by $\mathfrak{p} \mapsto \mathfrak{p}_0$). By (unr) and (ord), we may identify

$$U(\mathbb{Q}_p) = \prod_{\mathfrak{p} \in \Sigma_{sp}} GL_n(F_\mathfrak{p}) \times \prod_{\mathfrak{p} \in \Sigma_i} U_0(F_{0,\mathfrak{p}})$$

for the group U_0 in (8.30). Define a parabolic subgroup $P \subset GU(\mathbb{Q}_p)$ by the normalizer of

$$P_U = \left\{(x_\mathfrak{p})_{\mathfrak{p} \in \Sigma} \in U(\mathbb{Q}_p) | x_\mathfrak{p} = \left(\begin{smallmatrix} a & b \\ 0 & c \end{smallmatrix}\right) \text{ with } a \in GL_{m(\mathfrak{p})}(F_\mathfrak{p})\right\}. \tag{8.31}$$

Thus if $\mathfrak{p} \in \Sigma_i$, the $\mathfrak{p}_0$-component of P is the upper triangular balanced parabolic subgroup of $GU_0(F_{0,\mathfrak{p}})$.

We assume that we are in Case C and that (GU, X) as in (C) of Example 7.6 yields a Shimura variety of PEL type. Then $GU(\mathbb{R}) \cong GSp_{2n}(\mathbb{R})^{I_F}$. In this case, we define P to be the Siegel parabolic subgroup P_n, because $GU(\mathbb{Q}_p) \cong \prod_{\mathfrak{p} \in \mathcal{S}} GSp_{2n}(F_\mathfrak{p})$.

We now return to the general case where GU can be of type A or C. Recall the p-adic valuation ord_p of the strict Henselization $\mathcal{W} \subset \overline{\mathbb{Q}}$ of $\mathbb{Z}_{(p)}$ coming from the embedding $i_p : \overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_p$. Let $v = v_{GU}$ be the discrete valuation of $\mathfrak{K}$ over ord_p such that the residue field of v is an algebraic extension of the function field of an irreducible component of the Igusa tower over the ordinary locus S_1 of $\mathfrak{M}^{(p)}_{GU/\kappa}$. Recall the constant field $\mathfrak{k} = \mathfrak{K} \cap \overline{\mathbb{Q}}$. Let

$$\mathcal{D} = \left\{\sigma \in \mathrm{Aut}(\mathfrak{K}/E) \;\middle|\; v_{GU} \circ \sigma = v_{GU} \text{ and } \sigma \in \tau(\mathcal{G}(GU, X))\right\},$$
$$\mathfrak{D} = \{\sigma \in \mathrm{Gal}(\mathfrak{k}/E) \mid ord_p \circ \sigma = ord_p\}.$$

Thus $\mathcal{D}$ is the decomposition (or monodromy) group of v_{GU} inside $\mathrm{Im}(\tau)$.

We can make explicit the valuation v_{GU} as follows. We first deal with the unitary case. We pick a CM point $x \in \mathfrak{M}^{(p)}_{GU}(W)$, and suppose that the

abelian scheme $A = A_x$ sitting over x is p-ordinary. Since x is a W-point, by the universality of $\mathfrak{M}^{(p)}_{GU}$, $A_{x/W}$ is an abelian scheme.

We have identified L with $\mathbb{Z}^g \times \mathbb{Z}^g$ to embed G_1 into $Sp(2g)$ by ϕ. Let L_p^{ord} be the first factor $\mathbb{Z}_p^g$ (which is a totally isotropic sublattice under the alternating form $\langle \cdot, \cdot \rangle$ in (L1–2) of Section 7.1.1). Since $A_{x/W}$ is ordinary, we can fix an ordinary level structure $\eta_p^{ord} : L_p^{ord} \otimes \mu_{p^\infty} \cong A_x[p^\infty]^\circ$, which induces by duality composed with the polarization $\eta_p^{et} : L_p^{et} \otimes (\mathbb{Q}_p/\mathbb{Z}_p) \cong A_x[p^\infty]^{et}$ for $L_p^{et} = L_p/L_p^{ord}$. Thus we have a lift of the point x to the Igusa tower $T_\infty = \mathrm{Isom}_{O_D}(L_p^{ord} \otimes \mu_{p^\infty}, \mathcal{A}^{ord}_{GU}[p^\infty]^\circ)$, which we again denote by x. Then via the Serre–Tate deformation theory, we computed in Section 8.2 the formal completion of $Sh_{K/W}$ along $x \in Sh_K(W)$ as the formal deformation space $\widehat{S}_{GU/W}$ of A_x. The result is as follows. Write $\mathcal{T}^{et}_{\mathfrak{p}}(A)$ ($\mathfrak{p} \in \mathcal{S}$) for the $\mathfrak{p}$-adic Tate module of the $\mathfrak{p}$-power torsion points of $A_x(\mathbb{F})$. By (8.20) and (8.21),

$$\widehat{S}_{GU}(R) = \prod_{\mathfrak{p} \in \Sigma} \mathrm{Hom}(\varepsilon \cdot \mathcal{T}^{et}_{\mathfrak{p}}(A) \otimes_O \varepsilon \cdot \mathcal{T}^{et}_{\mathfrak{p}^c}(A), \widehat{\mathbb{G}}_m(R)) \xrightarrow[\eta_p^{et}]{\sim} \mathrm{Hom}(\varepsilon \cdot L^{et}_{\Sigma} \otimes_O \varepsilon \cdot L^{et}_{\Sigma^c}, \widehat{\mathbb{G}}_m(R))$$

for p-adic W-algebras R, where we have written $L^{et}_{\Sigma} = (L_p/L_p^{ord}) \otimes_O O_\Sigma$ and $L^{et}_{\Sigma^c} = (L_p/L_p^{ord}) \otimes_O O_{\Sigma^c}$. Here $O_\Sigma = \prod_{\mathfrak{p} \in \Sigma} O_{\mathfrak{p}}$. Regarding A_x as a member $A_{\iota(x)}$ of the universal family over the (geometrically irreducible) Siegel modular variety $\mathfrak{M}^{(p)}_{/W}$, by (8.22), the formal completion $\widehat{S}$ along x in $Sh_{K'}(GSp(2g), \mathcal{X})(W)$ is given by

$$\widehat{S}(R) = \mathrm{Hom}\big(Sym^2(\mathcal{T}^{et}_p(A)), \widehat{\mathbb{G}}_m(R)\big) \xrightarrow[\eta_p]{\sim} \mathrm{Hom}(Sym^2(L^{et}_p), \widehat{\mathbb{G}}_m(R))$$

where the symbol Sym^2 indicates the symmetric square tensor over $\mathbb{Z}_p$ and $L^{et}_p = L_p/L_p^{ord}$. The morphism $i : \mathfrak{M}^{(p)}_G \to \mathfrak{M}^{(p)}$ restricted to $\widehat{S}_{GU}$ is the embedding induced by the projection $Sym^2(L^{et}_p) \twoheadrightarrow \varepsilon L^{et}_{\Sigma} \otimes_O \varepsilon L^{et}_{\Sigma^c}$, which associates a symmetric form $\Psi(x,y) = \psi(x,y) + \psi(y,x) \in \widehat{S}(R)$ with each O_D-bilinear form $\psi : \mathcal{T}^{et}_p(A) \times \mathcal{T}^{et}_p(A) \to \widehat{\mathbb{G}}_m(R)$ in $\widehat{S}_{GU}(R)$. Note that the O_D-linear form ψ is determined by its restriction to the subspace $\prod_{\mathfrak{p} \in \Sigma} \big(\varepsilon \cdot \mathcal{T}^{et}_{\mathfrak{p}}(A) \times \varepsilon \cdot \mathcal{T}^{et}_{\mathfrak{p}^c}(A)\big)$; so, $\psi \mapsto \Psi$ gives rise to the closed immersion of $\widehat{S}_{GU}$ into $\widehat{S}$. The restriction $f \mapsto f|_{\widehat{S}_{GU}}$ for $f \in \mathcal{O}_{\widehat{S}}$ gives rise to a surjective ring homomorphism $\pi : \mathcal{O}_{\widehat{S}} \twoheadrightarrow \mathcal{O}_{\widehat{S}_{GU}}$ with $\mathrm{Ker}(\pi) = \mathfrak{Q}\mathcal{O}_{\widehat{S}}$, which shows that the morphism i is a closed immersion over $\mathcal{V}$ of the ordinary locus of $Sh^{(p)}_G$ into that of the Siegel modular variety (as claimed in the previous subsection).

Writing $\mathcal{O}_{\widehat{S}_{GU}} = W[[t]]$ for variables $t = (t_{\mathfrak{p},ij})_{\mathfrak{p} \in \Sigma, i, j}$ for $1 \le i \le m(\mathfrak{p})$ and $1 \le j \le m(\mathfrak{p}^c)$, we can make $v = v_{GU}$ explicit. Write $\mathfrak{K}^{(p)}$ for the function field of the geometrically irreducible component $\mathfrak{M}^{(p)}_{GU}$ of $Sh^{(p)}_{GU}$. Recall the subring $\mathfrak{V}$ of the Siegel modular function field $\mathfrak{F}$ and its prime ideal $\mathfrak{q}$ (vanishing on

the image of the unitary Shimura variety in the Siegel modular variety) in Theorem 8.15. Expanding $f \in (\mathfrak{V}/\mathfrak{q}) \cap \mathfrak{K}^{(p)} \subset \mathcal{O}_{\widehat{S}_{GU}}$ into $\sum_\xi a_\xi t^\xi$, we have a valuation $v_{GU}(f) = \mathrm{Inf}_\xi\, ord_p(a_\xi)$. Then v_{GU} over $\mathfrak{K}$ is an extension of the above v_{GU} to $\mathfrak{K}^{(p)}$ (whose residue field gives rise to the function field of an irreducible component of the Igusa tower T_∞).

Recall the function field $\mathfrak{F}^{(p)}$ of the generic fiber of the geometrically irreducible component $\mathfrak{M}^{(p)}$ of $Sh^{(p)}_{GSp(2g)}$. Similarly as above, $i(x)$ in the Igusa tower over $Sh^{(p)}_{GSp(2g)}(W)$ gives a point in the irreducible component containing the infinity cusp (as we have already proven the irreducibility in this case: Theorem 6.27). Thus by Theorem 6.28, the original v defined for $\mathfrak{F}^{(p)}$ in the proof of Theorem 6.27 can be rewritten in the same manner as $v_{GSp(2g)}(f) = \mathrm{Inf}_\xi\, ord_p(c_\xi)$, where we have expanded $f \in \mathfrak{F}^{(p)}$ into $\sum_\xi c_\xi T^\xi$ for the variables T with $\mathcal{O}_{\widehat{S}} = W[[T]]$. This is because we can measure the divisibility by p of an algebraic function on a geometrically irreducible smooth scheme over $\mathcal{W}$ by the expansion at any closed point of the scheme with respect to the local coordinate over W. In our case, one point is given by the cusp ∞ and another given by $i(x)$ (since $i(x)$ and ∞ are sitting on the same geometrically irreducible component of the special fiber). In particular, choosing T suitably, $\widehat{S}_{GU}$ is a complete intersection defined by $T' = 0$ for a suitable subset $T' \subset T$ of variables in T. The generic transversality of $\mathfrak{M}^{(p)}_{GU/\mathcal{W}}$ and $\mathfrak{M}^{(p)}_{/\mathbb{F}}$ inside $\mathfrak{M}^{(p)}_{/\mathcal{W}}$ claimed in the previous subsection follows from this expression. By this fact, the transversality is indeed valid in the entire ordinary locus of $\mathfrak{M}^{(p)}_{/\mathcal{W}}$. This also shows that, choosing the extension of v_{GU} defined above for $\mathfrak{K}^{(p)}$ to $\mathfrak{K}$ compatibly with v on $\mathfrak{F}$, the localization of $\mathfrak{V}/\mathfrak{q}$ at $\mathfrak{m}_{GU}$ is the valuation ring of the extension v_{GU} to $\mathfrak{K}$. In other words, the decomposition group $\mathcal{D}$ for v_{GU} is given by the stabilizer $\mathcal{D}_{GU,v}$ of $\mathfrak{m}_{GU}$ studied in Theorem 8.15.

The symplectic case is similar. In this case, we have by (8.22),

$$\widehat{S}_{GU}(R) = \prod_{\mathfrak{p}} \mathrm{Hom}\left(Sym^2(\varepsilon \cdot \mathcal{T}^{et}_{\mathfrak{p}}(A)), \widehat{\mathbb{G}}_m(R)\right),$$

where $\mathfrak{p}$ runs over all p-adic places of the base (totally real) field F and the symbol Sym^2 indicates the symmetric square tensor over $O_{\mathfrak{p}}$. The same argument as above proves that the decomposition group for v_{GU} is given by the stabilizer of $\mathfrak{p}_{GU}$ in Theorem 8.15. We get for groups of type A and C:

Theorem 8.16 *Suppose* (unr) *and* (ord). *Then the group $\mathcal{D}$ is the image (under τ) of a conjugate in $\mathcal{G}(GU, X)$ under* $\mathrm{Aut}(\mathfrak{M}_{GU})$ *of*

$$\mathcal{P} = \left\{ x \in \left(P \times GU(\mathbb{A}^{(p)})\right) \bigcap \mathcal{G}(GU, X) \middle| \sigma(x) \in \mathfrak{D} \right\},$$

where σ is as in Theorem 7.7 (2) *and P is as in* (8.31) *in the unitary case and is the Siegel parabolic subgroup in the symplectic case.*

Proof. This directly follows from Theorem 6.28 and Theorem 8.15, because the valuation we created, v_{GU}, is given by $v = v_{GSp(2g)}$ on $\mathfrak{F}$ induced from the q-expansion by the choice of our embedding $\phi : GU \to GSp(2g)$. For general v_{GU} not necessarily related to Siegel modular q-expansions, its decomposition group in $\mathrm{Aut}(\mathfrak{K})$ is the conjugate of the one specifically chosen. Since the image of $\mathcal{G}(G, X)$ in $\mathrm{Aut}(\mathfrak{K})$ is a normal subgroup of finite index, the conjugated image of $\mathcal{P}$ remains in that of $\mathcal{G}(G, X)$. □

We consider the universal abelian scheme $\mathcal{A}^{ord}_{GU/S}$ over $S_1 = Sh^{(p)}_{GU}\left[\frac{1}{H}\right] \otimes_W \mathbb{F}$ for the Hasse invariant H. Then we recall the Igusa tower over S_1 as follows. We write $\mathcal{A}^{ord}_{GU}[p^m]^{et}_{/S_1}$ for the maximal étale quotient of $\mathcal{A}^{ord}_{GU}[p^m]$. Then we define a constant group scheme C_m by

$$C_{m/S_1} = \begin{cases} \prod_{\mathfrak{p}\in\mathcal{S}} (O_{\mathfrak{p}}/p^m O_{\mathfrak{p}})^{m(\mathfrak{p})} & \text{in Case A} \\ (O_F/p^m O_F)^n & \text{in Case C.} \end{cases}$$

The Igusa tower is given by $T_{m/S_1} = \mathrm{Isom}_O\left(C_{m/S_1}, \varepsilon\left(\mathcal{A}^{ord}_{GU}[p^m]^{et}\right)_{/S_1}\right)$, where $GU(\mathbb{Q}_p) \cong GSp_{2n}(F \otimes_{\mathbb{Q}} \mathbb{Q}_p)$ in Case C.

Corollary 8.17 *Let M be the reductive part of the parabolic subgroup P_U in* (8.31). *Define $M_1 = M(\mathbb{Z}_p) \cap G_1(\mathbb{Z}_p)$ (so, $M_1 = M(\mathbb{Z}_p)$ in Case C). Then under the hypotheses of* Theorem 8.16, *each geometrically connected component T°_∞ of the generalized Igusa tower T_∞ over a geometrically irreducible component S°_1 of S_1 has Galois group over S°_1 isomorphic to M_1. In particular, $T_\infty \times_{S_1} S^\circ_1$ is irreducible if GU is of type C, and if $Sh^{(p)}(GU, X_{GU})$ is not projective, each geometrically irreducible component of $T_\infty \times_{S_1} S^\circ_1$ contains at least one cusp.*

Except for the ramified primes p in $F/\mathbb{Q}$ in the Hilbert modular case, this corollary includes all earlier results on mod p irreducibility under (unr) as special cases and also can be applied to projective Shimura varieties.

Proof. We take a characteristic 0 lift E of the Hasse invariant and consider $S = Sh^{(p)}_{GU}[\frac{1}{E}]_{/\mathcal{W}}$. We put $S_m = S \otimes_W W_m$ and $S_\infty = \varprojlim_m S_m$, which is independent of the choice of E. We recall that $\mathfrak{M}^{(p)}_{GU/\mathcal{W}} = \varprojlim_S \mathfrak{M}_{S/\mathcal{W}}$ for $\mathfrak{M}_S$ with $\mathfrak{M}_S(\mathbb{C}) = \Gamma_S \backslash X^+$, where $\Gamma_S = SU(\mathbb{Q}) \cap S \cdot GU(\mathbb{R})$ and S runs over all open compact subgroups $S \subset GU(\mathbb{A}^{(\infty)})$ maximal at p. We may assume that $S^\circ_1 = \mathfrak{M}^{(p)}_{GU}[\frac{1}{E}] \otimes_{\mathcal{W}} \mathbb{F}$ by the existence of a smooth projective compactification of $\mathfrak{M}^{(p)}_{GU/W}$ ([F]). Each element $g \in \mathcal{P}$ acts on S_m, because it preserves the ordinary locus. Suppose that GU is of type A. By (5.34), $\det(\underline{\omega}^\sigma) \cong \det(\underline{\omega}^{\sigma c})$ on $\Gamma_S \backslash X^+$. Note that $\mathfrak{M}_{S/W}$ is strongly quasi-projective (by the existence of the projective smooth toroidal compactification; see [F]). Then by [NMD] Theorem 5 in Section 8.2, $\mathrm{Pic}_{\mathfrak{M}_{S/W}}$ is a quasi-projective scheme. Since $\underline{\omega}^\sigma, \underline{\omega}^{\sigma c} \in \mathrm{Pic}_{\mathfrak{M}_{S/W}}(W)$, we find that $\det(\underline{\omega}^\sigma) \cong \det(\underline{\omega}^{\sigma c}) \otimes \mathcal{L}$ over W up to $\mathcal{L} \in \mathrm{Pic}_W(W)$ which is trivial. Thus we get $\det(\underline{\omega}^\sigma)_{/W} \cong$

$\det(\underline{\omega}^{\sigma c})_{/W}$. Since over S_∞, we have $\det(\underline{\omega}^{\sigma})_{/W} = \det(\varepsilon\mathcal{T}(\mathcal{A}_{GU}[\mathfrak{p}^\infty]^{et})) \otimes_{\mathbb{Z}_p} \mathcal{O}_{S_\infty}$ (where $\det(\varepsilon\mathcal{T}(\mathcal{A}_{GU}[\mathfrak{p}^\infty]^{et})) = \bigwedge_O^{m(\mathfrak{p})} \det(\varepsilon\mathcal{T}(\mathcal{A}_{GU}[\mathfrak{p}^\infty]^{et}))$ for the ideal $\mathfrak{p}$ of F induced by $i_p \circ \sigma$), we find from the faithful flatness of $\mathcal{O}_{S_\infty}$ over $\mathbb{Z}_p$ that $\det(\varepsilon\mathcal{A}_{GU}[\mathfrak{p}^n]^{et}) \cong \det(\varepsilon\mathcal{A}_{GU}[\mathfrak{p}^{cn}]^{et})$ over S_m° for all m and n. This implies that for $g \in M(\mathbb{Z}_p)$ to be in $\mathcal{P}$, it has to belong to M_1 which is given by

$$M_1 = \begin{cases} \left\{ (g_\mathfrak{p})_\mathfrak{p} \in \prod_\mathfrak{p} GL_{m(\mathfrak{p})}(O_\mathfrak{p}) \middle| \det(g_\mathfrak{p}) = \det(g_{\mathfrak{p}^c}) \ \forall \mathfrak{p} \in \mathcal{S} \right\} & \text{in Case A,} \\ \prod_{\mathfrak{p} \in \mathcal{S}} GL_n(O_\mathfrak{p}) & \text{in Case C.} \end{cases}$$

In Case C, there is no restriction. The fact $M_1 = M(\mathbb{Z}_p) \cap \mathcal{P}$ can also be checked by computing $\mathcal{G}(GU, X)$ explicitly. Since $M_1 \subset \mathcal{P}$, $M_1 \hookrightarrow \mathcal{D}$ by Theorem 8.16, and this shows $M_1 = \mathrm{Gal}(T_\infty^\circ / S_\infty^\circ)$ for the irreducible component $T_{\infty/S_1^\circ}^\circ$ of T_∞ whose decomposition group is $\mathcal{D}$. In Case C, the Galois group is full, and hence $T_\infty \times_{S_1} S_1^\circ$ is irreducible. In Case A, we have $\pi_0(T'_\infty) = M(\mathbb{Z}_p)/M_1$ for $T'_\infty = T_\infty \times_{S_1} S_1^\circ$, because $M(\mathbb{Z}_p)$ acts transitively on $\pi_0(T'_\infty)$. Then the ordinary locus of the minimal compactification T_∞^* of T'_∞ has to have an irreducible component containing a cusp, because T'_∞ is a Galois covering of S_1° whose compactification has a cusp. Since $M(\mathbb{Z}_p)$ acts transitively on the irreducible components of T'_∞, every component contains a cusp. □

By the above result, in the symplectic case, the q-expansion principle holds as usual (as already stated for the Hilbert modular and the Siegel modular cases). In the unitary case, to determine a mod p (or p-adic) automorphic form by q-expansion, one needs an expansion at one cusp for each geometrically irreducible component of the Shimura variety before the reduction modulo p. We leave for the reader a precise formulation of this fact.

References

Books

[AAF] G. Shimura, *Arithmeticity in the Theory of Automorphic Forms*, Mathematical Surveys and Monographs **82**, American Mathematical Society, Providence, RI, 2000.

[AAG] S. S. Gelbart, *Automorphic Forms on Adele Groups*, Annals of Math. Studies **83**, Princeton University Press, Princeton, NJ, 1975.

[AAQ] M.-F. Vignéras, *Arithmétique des algèbres de quaternions*, Lecture Notes in Mathematics **800**, Springer, New York, 1980.

[ABV] D. Mumford, *Abelian Varieties*, TIFR Studies in Mathematics, Oxford University Press, New York, 1994.

[ACM] G. Shimura, *Abelian Varieties with Complex Multiplication and Modular Functions*, Princeton University Press, Princeton, NJ, 1998.

[AGD] A. Borel and G. D. Mostow, editors, *Algebraic Groups and Discontinuous Subgroups*, Proc. Symp. Pure Math. IX, American Mathematical Society, Providence, RI, 1966.

[ALF] K. Iwasawa, *Algebraic Functions*, Translation from the 1973 Japanese edition by Goro Kato. Translations of Mathematical Monographs, **118**, American Mathematical Society, Providence, RI, 1993.

[ALG] R. Hartshorne, *Algebraic Geometry*, Graduate Texts in Mathematics **52**, Springer, New York, 1977.

[AME] N. M. Katz and B. Mazur, *Arithmetic Moduli of Elliptic Curves*, Annals of Math. Studies **108**, Princeton University Press, Princeton, NJ, 1985.

[ARE] J.-P. Serre, *Abelian ℓ-Adic Representations and Elliptic Curves*, Benjamin, New York, 1968.

[ARG] G. Cornell and J. H. Silverman, editors, *Arithmetic Geometry*, Springer, New York, 1986.

[BAL] N. Bourbaki, *Algébre*, Hermann, Paris, 1958.

[BCM] N. Bourbaki, *Algbre Commutative*, Hermann, Paris, 1961–1965.

[BLI] N. Bourbaki, *Groupes et Algébres de Lie*, Hermann, Paris, 1972–1985.

[BNT] A. Weil, *Basic Number Theory*, Springer, New York, 1974.

[CBT] W. Messing, *The Crystals Associated to Barsotti–Tate Groups; With Applications to Abelian Schemes*, Lecture Notes in Mathematics **264**, New York, Springer, 1972.

[CDR] A. Borel and N. Wallach, *Continuous Cohomology, Discrete Subgroups, and Representations of Reductive Groups*, Annals of Math. Studies **94**, Princeton University Press, Princeton, NJ, 1980.

[CFT] E. Artin and J. Tate, *Class Field Theory*, Benjamin, New York, 1968.

[CGP] K. S. Brown, *Cohomology of Groups*, Graduate Texts in Mathematics **87**, Springer, New York, 1982.

[CPS] G. Shimura, *Collected Papers*, I, II, III, IV, Springer, New York, 2002.

[CRT] H. Matsumura, *Commutative Ring Theory*, Cambridge Studies in Advanced Mathematics **8**, Cambridge Univ. Press, New York, 1986.

[CSM] C.-L. Chai, *Compactification of Siegel Moduli Schemes*, LMS Lecture Note Series **107**, Cambridge University Press, New York, 1985.

[DAV] G. Faltings and C.-L. Chai, *Degeneration of Abelian Varieties*, Springer, New York, 1990.

[DGH] J. Tilouine, *Deformations of Galois Representations and Hecke Algebras* (published for the Mehta Research Institute of Mathematics and Mathematical Physics, Allahabad), Narosa Publishing House, New Delhi, 1996.

[ECH] J. S. Milne, Étale Cohomology, Princeton University Press, Princeton, NJ, 1980.

[EGA] A. Grothendieck and J. Dieudonné, *Eléments de Géométrie Algébrique*, Publications IHES **4** (1960), **8** (1961), **11** (1961), **17** (1963), **20** (1964), **24** (1965), **28** (1966), **32** (1967).

[EPE] G. Shimura, *Euler Products and Eisenstein Series*, CBMS Regional Conference Series **93**, American Mathematical Society, Providence, RI, 1997.

[FAN] D. Ramakrishnan and R. J. Valenza, *Fourier Analysis on Number Fields*, Graduate Texts in Mathematics **186**, Springer, New York, 1999.

[FGA] A. Grothendieck, *Fondements de la Géométrie Algébrique*, Séminaire Bourbaki exp. no.149 (1956/57), 182 (1958/59), 190 (1959/60), 195 (1959/60), 212 (1960/61), 221 (1960/61), 232 (1961/62), Benjamin, New York, 1966.

[GAN] M. Lazard, *Groupes Analytiques p–Adiques*, Publications IHES **26**, 1965.

[GCH] J.-P. Serre, *Galois Cohomology*, in Monographs in Mathematics, Springer, New York, 2002.

[GCS] M. Harris and R. Taylor, *The Geometry and Cohomology of Some Simple Shimura Varieties*, Annals of Math. Studies **151**, Princeton University Press, Princeton, NJ, 2001.

[GIT] D. Mumford, *Geometric Invariant Theory*, Ergebnisse **34**, Springer, New York, 1965.

[GME] H. Hida, *Geometric Modular Forms and Elliptic Curves*, World Scientific, Singapore, 2000.

[HAL] P. J. Hilton and U. Stammback, *A Course in Homological Algebra*, Graduate Texts in Mathematics **4**, Springer, New York, 1970.

[HMS] P. Deligne, J. S. Milne, A. Ogus and K.-y. Shih, *Hodge Cycles, Motives, and Shimura Varieties*, Lecture Notes in Mathematics **900**, Springer, New York, 1982.

[IAT] G. Shimura, *Introduction to the Arithmetic Theory of Automorphic Functions*, Princeton University Press, Princeton, NJ, and Iwanami Shoten, Tokyo, 1971.

[ICF] L. C. Washington, *Introduction to Cyclotomic Fields*, Graduate Text in Mathematics, **83**, Springer, New York, 1980.

[LAP] H. Yoshida, *Lectures on Absolute CM Period*, Mathematical Surveys and Monographs **106**, American Mathematical Society, Providence, RI, 2003.

[LFE] H. Hida, *Elementary Theory of L-Functions and Eisenstein Series*, LMSST **26**, Cambridge University Press, Cambridge, England, 1993.

[MFG] H. Hida, *Modular Forms and Galois Cohomology*, Cambridge Studies in Advanced Mathematics **69**, Cambridge University Press, Cambridge, England, 2000.

[MFM] T. Miyake, *Modular Forms*, Springer, New York-Tokyo, 1989.

[MTV] U. Jannsen, S. Kleiman and J.-P. Serre, *Motives*, Proc. Symp. Pure Math. **55** Part 1 and 2, American Mathematical Society, Providence, RI, 1994.

[NAZ] N. Koblitz, *p-Adic Numbers, p-Adic Analysis, and Zeta Functions*, Graduate Text in Mathematics, **58**, Springer, New York, 1977.

[NMD] S. Bosch, W. Lütkebohmert, and M. Raynaud, *Néron Models*, Springer, New York, 1990.

[RAG] J. C. Jantzen, *Representations of Algebraic Groups*, Academic Press, Orlando, FL, 1987.

[RTP] G. Faltings and G. Wüstholtz et al, *Rational Points*, Aspects of Mathematics **E6**, Friedr. Vieweg & Sohn, Braunschweig, 1992.

[SFT] G. E. Breadon, *Sheaf Theory*, McGraw-Hill, New York, 1967.

[SGA] A. Grothendieck, *Revetements Étale et Groupe Fondamental*, Séminaire de geometrie algébrique, Lecture Notes in Mathematics **224**, Springer, New York, 1971.

[SGL] H. Hida, *On the Search of Genuine p-Adic Modular L-Functions for $GL(n)$*, Mem. SMF **67**, 1996.

[TEB] G. Kempf, F. Knudson, D. Mumford and B. Saint-Donat, *Toroidal Embeddings I*, Lecture Notes in Mathematics **339**, Springer, New York, 1973.

[TFG] A Grothendieck and J. P. Murre, *The Tame Fundamental Group of a Formal Neighborhood of a Divisor with Normal Crossings on a Scheme*, Lecture Notes in Mathematics **208**, Springer, New York, 1971.

Articles

[BeZ] I. N. Bernstein and A. V. Zelevinsky, Induced representations of reductive $\mathfrak{p}$–adic groups, I, Ann. Sci. Ec. Norm. Sup. 4th series **10** (1977), 441–472.

[Bl] D. Blasius, Elliptic curves, Hilbert modular forms, and the Hodge conjecture, in *Contributions to Automorphic Forms, Geometry, and Number Theory*, a supplement of the American Journal of Mathematics, Johns Hopkins University Press, Baltimore, MD, 2004.

[BlR] D. Blasius and J. D. Rogawski, Motives for Hilbert modular forms, Inventiones Math. **114** (1993), 55–87.

[BlR1] D. Blasius and J. D. Rogawski, Zeta functions of Shimura varieties, In "Motive": [MTV], Proc. Symp. Pure Math. **55** (1994) Part 2, 525–571.

[Bo] A. Borel, Reduction theory of arithmetic groups, Proc. Symp. Pure Math. **9** (1966), 20–25 (Œuvres II 679–684, No.74).

[BoS] A. Borel and J. P. Serre, Corners and arithmetic groups, Comment. Math. Helv. **48** (1974), 436–491 (Borel's Œuvres III 244–299 No.98).

[Br] L. Breen, Tannakian categories, Proc. Symp. Pure Math. **55** (1994) Part I ([MTV] I), 337–376.

[BrCDT] C. Breuil, B. Conrad, F. Diamond, and R. Taylor, On the modularity of elliptic curves over $\mathbb{Q}$ or Wild 3-adic exercises, Journal AMS **14** (2001), 843–939.

[BuDST] K. Buzzard, M. Dickinson, N. Shepherd-Barron, and R. Taylor, On icosahedral Artin representations, Duke Math. J. **109** (2001), 283–318.

[BuT] K. Buzzard and R. Taylor, Companion forms and weight one forms, Ann. of Math. **149** (1999), 905–919.

[C] C.-L. Chai, Arithmetic minimal compactification of the Hilbert-Blumenthal moduli spaces, Appendix to [Wi1], Ann. of Math. **131** (1990), 541–554.

[Co] R. F. Coleman, p-adic Banach spaces and families of modular forms, Inventiones Math. **127** (1997), 417–479.

[D] P. Deligne, Variété abeliennes ordinaires sur un corps fini, Inventiones Math. **8** (1969), 238–243.

[D1] P. Deligne, Travaux de Shimura, Sem. Bourbaki, Exp. 389, Lecture Notes in Math. **244** (1971), 123–165.

[D2] P. Deligne, Variétés de Shimura: Interprétation modulaire, et techniques de construction de modéles canoniques, Proc. Symp. Pure Math. **33.2** (1979), 247–290.

[D3] P. Deligne, A quoi servent les motifs?, Proc. Symp. Pure Math. **55** (1994) Part I ([MTV] I), 143–161

[DM] P. Deligne and D. Mumford, The irreducibility of the space of curves of given genus, Publ. I.H.E.S. **36** (1969), 75–109.

[DR] P. Deligne and K. A. Ribet, Values of abelian L-functions at negative integers over totally real fields, Inventiones Math. **59** (1980), 227–286.

[DT] M. Dimitrov and J. Tilouine, Variété de Hilbert et arithmétique des formes modulaires de Hilbert pour $\Gamma_1(\mathfrak{c}, N)$, to appear in *Geometric Aspects of Dwork's Theory, A Volume in Memory of Bernard Dwork* (edited by A. Adolphson, F. Baldassarri, P. Berthelot, N. Katz, and F. Loeser), Walter de Gruyter, Berlin, 2004.

[Du] M. Deuring, Die Typen der Multiplikatorenringe elliptischer Funktionenkörper, Abhandlungen Math. Sem. Hansischen Universität **14** (1941), 197–272.

[F] K. Fujiwara, Arithmetic compactifications of Shimura varieties, Master's Thesis (University of Tokyo), 1989.

[F1] K. Fujiwara, Deformation rings and Hecke algebras in totally real case, preprint, 1999.

[GM] F. Q. Gouvêa and B. Mazur, Families of modular forms, Math. Comp. **58** (1992), 793–805.

[G] B. B. Gordon, Canonical models of Picard modular surfaces, CRM Publications **13** (1992), 1–29.

[Gu] F. Q. Gouvêa, Non-ordinary primes: a story. Experiment. Math. **6** (1997), 195–205.

[H79] H. Hida, On abelian varieties with complex multiplication as factors of abelian variety attached to Hilbert modular forms (Master's Thesis at Kyoto University, 1977), Japan. J. Math. **5** (1979), 157–208.

[H81] H. Hida, On abelian varieties with complex multiplication as factors of the Jacobians of Shimura curves (Doctor's Thesis at Kyoto University, 1980), Amer. J. Math. **103** (1981), 727–776.

[H85] H. Hida, A p–adic measure attached to the zeta functions associated with two elliptic modular forms. I, Inventiones Math. **79** (1985), 159–195.

[H86a] H. Hida, Iwasawa modules attached to congruences of cusp forms, Ann. Sci. Ec. Norm. Sup. 4th series **19** (1986), 231–273.

[H86b] H. Hida, Galois representations into $GL_2(\mathbb{Z}_p[[X]])$ attached to ordinary cusp forms, Inventiones Math. **85** (1986), 545–613.

[H88] H. Hida, On p-adic Hecke algebras for GL_2 over totally real fields, Ann. of Math. **128** (1988), 295–384.

[H89a] H. Hida, On nearly ordinary Hecke algebras for $GL(2)$ over totally real fields, Adv. Studies in Pure Math. **17** (1989), 139–169.

[H89b] H. Hida, Nearly ordinary Hecke algebras and Galois representations of several variables, Proc. JAMI Inaugural Conference, Supplement to Amer. J. Math. (1989), 115–134.

[H91] H. Hida, On p-adic L-functions of $GL(2) \times GL(2)$ over totally real fields, Ann. Inst. Fourier **41** (1991), 311–391.

[H94] H. Hida, On the critical values of L–functions of $GL(2)$ and $GL(2) \times GL(2)$, Duke Math. J. **74** (1994), 431–529.

[H95] H. Hida, Control theorems of p–ordinary cohomology groups for $SL(n)$, Bull. SMF **123** (1995), 425–475.

[H98] H. Hida, Automorphic induction and Leopoldt type conjectures for $GL(n)$, Asian J. Math. **2** (1998), 667–710.

[H02] H. Hida, Control theorems of coherent sheaves on Shimura varieties of PEL–type, Journal of the Inst. of Math. Jussieu, 2002 **1**, 1–76 (preprint downloadable at www.math.ucla.edu/~hida).

[H03a] H. Hida, p–Adic automorphic forms on reductive groups, to appear in The Proceedings of the Automorphic Semester of the Year 2000 at l'Institut Henri Poincaré (preprint downloadable at www.math.ucla.edu/~hida).

[H03b] H. Hida, Automorphism groups of Shimura varieties of PEL type, preprint, 2003 (preprint downloadable at www.math.ucla.edu/~hida).

[HM] H. Hida and Y. Maeda, Non-abelian base-change for totally real fields, Special Issue of Pacific J. Math. in memory of Olga Taussky Todd, 189–217, 1997.

[HT] H. Hida and J. Tilouine, Anticyclotomic Katz p–adic L–functions and congruence modules, Ann. Sci. Ec. Norm. Sup. 4th series **26** (1993), 189–259.

[Hl] D. Hilbert, Mathematical problems, Bull. Amer. Math. Soc. **37** (2000), 407–436.

[Hz] A. Hurwitz, Über die Entwicklungkoeffizienten der lemniskatischen Funktionen, Math. Ann. **51** (1899), 196–226 (Werke II, 342–373).

[I] J. Igusa, Kroneckerian model of fields of elliptic modular functions, Amer. J. Math. **81** (1959), 561–577.

[K] N. M. Katz, p–adic properties of modular schemes and modular forms, Lecture notes in Math. **350** (1973), 70–189.

[K1] N. M. Katz, Higher congruences between modular forms, Annals of Math. **101** (1975), 332–367.

[K2] N. M. Katz, Serre–Tate local moduli, In "Surfaces Algébriques," Lecture Notes in Math. **868** (1978), 138–202

[K3] N. M. Katz, p-adic L-functions for CM fields, Inventiones Math. **49** (1978), 199–297.

[Kn] M. Kneser, Strong approximation, in [AGD]: Proc. Symp. Pure Math. **IX** (1966), 187–196.

[Kn1] M. Kneser, Hasse principle for H^1 of simply connected groups, in [AGD]: Proc. Symp. Pure Math. **IX** (1966), 159–163.

[Ko] R. Kottwitz, Points on Shimura varieties over finite fields, J. Amer. Math. Soc. **5** (1992), 373–444.

[L] M. J. Larsen, Arithmetic compactification of some Shimura surfaces, CRM Publications **13** (1992), 31–45.

[LaR] R. P. Langlands and M. Rapoport, Shimuravarietäten und Gerben, J. Reine. Angew. Math. **378** (1987), 113–220.

[MaM] Y. Matsushima and S. Murakami, On vector bundle valued harmonic forms on symmetric spaces, Ann. of Math. **78** (1963), 365–416.

[Mi] J. Milne, Canonical models of (mixed) Shimura varieties and automorphic vector bundles, Perspective Math. **10** (1990), 283–414.

[Mi1] J. Milne, The points on a Shimura variety modulo a prime of good reduction, CRM Publications **13** (1992) ([ZFP]), 151–253

[Mi2] J. Milne, Shimura varieties and motives, Proc. Symp. Pure Math. **55** (1994) ([MTV]) Part 2, 447–523.

[MiS] J. Milne and K-y. Shih, Automorphism groups of Shimura varieties and reciprocity laws, Amer. J. Math. **103** (1981), 911–935.

[Mt] T. Miyake, On automorphism groups of the fields of automorphic functions, Ann. of Math. **95** (1972), 243–252.

[Mu] D. Mumford, Biextensions of formal groups, in *Algebraic Geometry*, Oxford Univ. Press, Oxford (1969), 307–322.

[Mu1] D. Mumford, An analytic construction of degenerating abelian varieties over complete rings, Compositio Math. **24** (1972), 239–272.

[P] A. A. Panchishkin, On the Siegel–Eisenstein measure and its applications, in "Proceedings of the Conference on p-adic Aspects of the Theory of Automorphic Representations (Jerusalem, 1998)". Israel J. Math. **120** (2000), part B, 467–509.

[Ra] M. Rapoport, Compactifications de l'espace de modules de Hilbert–Blumenthal, Comp. Math. **36** (1978), 255–335.

[Ri] K. A. Ribet, P–adic interpolation via Hilbert modular forms, Proc. Symp. Pure Math. **29** (1975), 581–592

[SBT] I. Shepherd-Barron and R. Taylor, Mod 2 and mod 5 icosahedral representations, J. AMS 10 (1997), 283–298.

[Sc] A. J. Scholl, Motives for modular forms, Inventiones Math. **100** (1990), 419–430.

[Se] J.-P, Serre, Géometrie algébrique et géometrie analytique, Annales Inst. Fourier **6** (1955), 1–42 (Œuvres I 402–443, No. 32).

[Se1] J.-P. Serre, Le probleme des groups de congruence pour SL_2, Ann. of Math. **92** (1970), 487–527 (Œuvres II 487–527, No. 86).

[Se2] J.-P, Serre, Propriétés galoisiennes des points d'ordre fini des courbes elliptiques, Inventiones Math. **15** (1972), 259–331 (Œuvres III 1–73, No. 94).

[Se3] J.-P. Serre, Formes modulaires et fonctions zêta p-adiques, Lecture notes in Math. **350** (1973), 191–268 (Œuvres III 95–172, No. 97).

[Se4] J.-P. Serre, Quelques applications du théorème de densité de Chebotarev, Publ. IHES. **54** (1981), 323–401 (Œuvres III 563–641, No. 125).

[Sh] G. Shimura, On analytic families of polarized abelian varieties and automorphic functions, Ann. of Math. **78** (1963), 149–192 ([63b] in [CPS] I).

[Sh1] G. Shimura, On modular correspondences for $Sp(N, \mathbb{Z})$ and their congruence relations, Proc. NAS **49** (1963), 824–828 ([63d] in [CPS] I).

[Sh2] G. Shimura, A reciprocity law in non-solvable extensions, J. Reine Angew. Math. 221 (1966), 209–220 ([66a] in [CPS] I).

[Sh3] G. Shimura, Moduli and fibre system of abelian varieties, Ann. of Math. **83** (1966), 294–338 ([66b] in [CPS] I).

[Sh4] G. Shimura, Construction of class fields and zeta functions of algebraic curves, Ann. of Math. **85** (1967), 58–159 ([67b] in [CPS] II).

[Sh5] G. Shimura, On canonical models of arithmetic quotients of bounded symmetric domains, Ann. of Math. **91** (1970), 144–222; II, **92** (1970), 528–549 ([70a–b] in [CPS] II).

[Sh6] G. Shimura, On some arithmetic properties of modular forms of one and several variables, Ann. of Math. **102** (1975), 491–515 ([75c] in [CPS] II).

[Sh7] G. Shimura, The special values of the zeta functions associated with cusp forms, Comm. Pure and Appl. Math. **29** (1976), 783–804 ([76b] in [CPS] II).

[Sh8] G. Shimura, On certain zeta functions attached to two Hilbert modular forms: I. The case of Hecke characters, II. The case of automorphic forms on a quaternion algebra, I: Ann. of Math. **114** (1981), 127–164; II: ibid. 569–607 ([81b–c] in [CPS] III).

[Si] K.-Y. Shih, Existence of certain canonical models, Duke Math. J. **45** (1978), 63–66.

[SkW] C. M. Skinner and A.J. Wiles, Residually reducible representations and modular forms. Publ. IHES. **89** (1999), 5–126.

[Sp] T. A. Springer, Galois cohomology of linear algebraic groups, in [AGD], Proc. Symp. Pure Math. **IX** (1966), 149–158.

[T] J. Tate, p-divisible groups, Proc. Conf. on local filds, Driebergen 1966, Springer 1967, 158–183.

[T1] J. Tate, Endomorphisms of abelian varieties over finite fields, Inventiones Math. **2** (1966), 134–144

[T2] J. Tate, Conjectures on algebraic cycles in ℓ-adic cohomology, in [MTV] Part 1, Proc. Symp. Pure Math. **55** (1994), 71–82.

[Ta] R. Taylor, On Galois representations associated to Hilbert modular forms, Inventiones Math. **98** (1989), 265–280.

[Ta1] R. Taylor, Icosahedral Galois representations, Pacific J. Math. (1997), Olga Tauski-Todd Memorial Issue 337–347.

[Ta2] R. Taylor, On icosahedral Artin representations II, Amer. J. Math. **125** (2003), 549–566.

[TaW] R. Taylor and A. Wiles, Ring theoretic properties of certain Hecke algebras, Ann. of Math. **141** (1995), 553–572.

[TiU] J. Tilouine and E. Urban, Several-variable p-adic families of Siegel–Hilbert cusp eigensystems and their Galois representations, Ann. Scient. Éc. Norm. Sup. 4th series **32** (1999), 499–574.

[Tt] J. Tits, Classification of algebraic semi-simple groups, Proc. Symp. Pure Math. **9** (1966), 33–62.

[Tt1] J. Tits, Reductive groups over local fields, in "Automorphic forms, representations and L-functions", Proc. Symp. Pure Math. **33** Part 1, (1979), 29–69.

[Wa] D. Wan, Dimension variation of classical and p-adic modular forms, Inventiones Math. **133** (1998), 449–463.

[We] A. Weil, Algebras with involutions and the classical groups, J. Indian Math. Soc. **24** (1960), 589–623 (Œuvres II 413–447, [1960b]).

[Wi] A. Wiles, On ordinary Λ-adic representations associated to modular forms, Inventiones Math. **94** (1988), 529–573.

[Wi1] A. Wiles, The Iwasawa conjecture for totally real fields, Ann. of Math. **131** (1990), 493–540.

[Wi2] A. Wiles, Modular elliptic curves and Fermat's last theorem, Ann. of Math. **141** (1995), 443–551.

[Z] J. G. Zarhin, Isogenies of abelian varieties over fields of finite characteristics, Math. USSR Sbornik **24** (1974), 451–461.

[Z1] J. G. Zarhin, Endomorphisms of abelian varieties over fields of finite characteristics, Math. USSR Izvestija **9 (2)** (1975), 255–260.

[Zi] T. Zink, Über die Schlechte Reduktion einiger Shimuramanifaltigkeiten, Compositio Math. **45** (1981), 15–107.

Symbol Index

$\langle \cdot, \cdot \rangle_N$, 103
$\langle z \rangle$, 180
$\langle \mathfrak{z} \rangle$, 125, 173

α^*, 167
$\Delta(E, \omega)$, 52
$\Delta_0(\mathfrak{N})$, 165
$\Delta_0(p^\infty \mathfrak{N})$, 168
Δ_H, 235
Δ_N, 235
ε_+, 169, 170
ε_Δ^v, 170
ε_S^v, 170
$\widehat{\Gamma}_1(\mathfrak{N})$, $\widehat{\Gamma}_1^1(\mathfrak{N})$, $\widehat{\Gamma}(\mathfrak{N})$, 108
μ_N, 71
$\mu_{\mathfrak{N}}$, 103
$\pi_0(T)$, 69
$\pi_1(S, s)$, $\pi_1(X, x)$, 133, 216
ρ_κ^C, 7
$\underline{\omega}_k$, 119
$\underline{\omega}_\varepsilon^\kappa$, $\underline{\omega}_{\kappa,\varepsilon}$, 171
$\omega(f)$, 246, 248

$\mathcal{B}$-ALG, $ALG_{/\mathcal{B}}$, 68
A_ε, 354
$\mathcal{C}(\mathfrak{a}, \mathfrak{b}, \phi_N)$, 110
$\mathfrak{c}(\lambda)$, 100
$\mathcal{D}(m, n)$, 239
$\mathcal{E}_{\Gamma_1^1(\mathfrak{N})}$, 103
$E_{X/S}^s$, 264
$\mathbf{e}(z)$, $\mathbf{e}(x)$, 24, 103
$ET_{/X}$, 216
$F_+^\times$, 100

$\mathfrak{F}$, $\mathfrak{F}^{(p)}$, 365
$G_k(\Gamma_1(N); R)$, 78, 86
$G_k(\Gamma_1(N); \mathbb{T}_p)$, $G_k^{ord}(\Gamma_1(N); R)$, 87
$G_\kappa(K, \varepsilon; R)$, $h_k(\mathfrak{N}, \chi; R)$, 177
$\mathfrak{H}$, 6
$h_\kappa^{n.ord}(K, \varepsilon; R)$, 189
$H_!^q$, 233
$\underline{\mathrm{Hom}}_S(P, P')$, $\underline{\mathrm{Isom}}_S(P, P')$, 297
$H(X, Y)_{/S}$, $H^{s,s'}(X, Y)_{/S}$, 263
i_∞, i_p, 12
Ig, $Ig_{S_0(p^r\mathfrak{N})}$, Ig_K, 164
$\mathrm{Jac}(V)$, 45
$\mathfrak{K}$, $\mathfrak{K}^{(p)}$, 365
$[KgK']$, 165
$M(\mathfrak{c}, \Gamma(N))$, 112
$\mathfrak{M}_{d,N}$, 277
$\mathfrak{M}_{GU}$, $\mathfrak{M}_{GU}^{(p)}$, 366, 367
$\mathrm{Pic}(V)$, $\mathrm{Pic}^0(V)$, 45
$\mathrm{Pic}_{X/S}$, 270
$\mathcal{R}$, $R(S, \Delta_S)$, 227, 228
$\mathrm{Res}_{R'/R} H_{/R'}$, 97
$SCH_{/\mathcal{B}}$, $\mathcal{B}$-SCH, 68
$S_1(\mathfrak{N})$, $S_1^1(\mathfrak{N})$, $S(\mathfrak{N})$, 108
$\mathbf{S}(\mathcal{E})$, 252
$\mathbb{S}_{/\mathbb{R}}$, 305
$Sh(G, X)$, $Sh^{(p)}$, 303, 309, 313, 325
$S_\kappa(K, \varepsilon; R)$, 177
$S_\kappa^{n.ord}(K, \varepsilon; R)$, 189
$\mathrm{Tate}_{\mathfrak{a},\mathfrak{b}}(q)$, 116
$T_v(y)$, $U_v(y)$, 174, 188
$\mathcal{V}_{cusp}(\widehat{\Gamma}_0(\mathfrak{N}))$, $V_{cusp}(\widehat{\Gamma}_0(\mathfrak{N}))$, 125
X_G, X_{GU}, 305, 306
$\mathfrak{z}$, 105

Statement Index

(Σ), 143

(Abel), 69
(A1–4), (A4′), 277, 278
(B1–2), 319
(C), 345
(C1–3), 211
(CT), 322
(D1–4), 321
(d1–2), 342
(det), 99, 268, 308
(E), 119, 338
(E1–3), 68
(ex0–3), 170
(F), 345
(F1–2), 212
(G0–2), 76
(G3), 79
(Hp1–2), 335
(Ism), 106
(L1–2), 304
(ord), 353, 369
(P1–3), 59, 350
(PC1–4), 110
(Pc1–4), 283
(pol), 134, 312, 315
(pos), 305
(QS), 219
(q-exp), 119
(rc1–2), 325
(rm1–4), 99
(S), 218
(S1–3), 218
(s1–3), 164
(SA1–3), 193
(SB), 203
(SB1–2), 201
(SC), 322
(sc), 305
(u1–4), 355
(unr), 304
(V1–4), 37

Subject Index

Abelian scheme, 98, 264
  of CM type, 4
  formal, 284
  ordinary, 101
  ordinary semi-abelian, 118
  polarization of, 99, 306
  Raynaud extension of, 284
  rigidity of endomorphisms, 264
  semi-abelian scheme, 116, 118, 265
  Weil pairing of, 103, 350
  with complex multiplication, 4
Abel's theorem, 56
Algebra
  double coset, 165, 227
  Hecke, 177, 189, 191
  Iwasawa, 8
  opposite, 351
Automorphic,
  factor, 6
  function field, 364
Automorphic form,
  central character, 171
  classical, 330, 335
  cusp form, 119
  false, 329, 335
  Λ-adic, 91
  p-adic analytic family of, 91
  true, 330, 335

Cartier duality, 71
Cases A, B, C, D, 316
Category,
  categorical quotient, 211
  direct product, 210
  direct sum, 210
  fibered category, 98, 99
  fibered product, 210
  fibered sum, 210
  final object, 210
  Galois, 212
  initial object, 210
  test object, 86
Character,
  arithmetic, 176
  Hecke, 176
  totally even, 177
CM field, 23, 154
CM point, 4, 154, 303
Cohomology,
  interior, 233
Compactification,
  minimal, 284
  toroidal, 284
Cone,
  decomposition, 110
  simplicial, 110
Congruence subgroup, 6
  cocompact, 6
Correspondence
  algebraic, 164
  Jacquet–Langlands and Shimizu, 202
Curve,
  Igusa, 82

Differential,
  holomorphic, 46
  of the first kind, 46

Divisor, 45
  algebraic equivalence of, 45
  degree of, 45
  linear equivalence of, 45
Duality,
  Cartier, 71

Elliptic curve, 21, 68
  j-invariant of, 22
  Tate, 78
  Weierstrass equation of, 52
Elliptic modular form,
  p-adic, 10
  q-expansion of, 63

Frobenius,
  absolute, 80
  element, 20
  relative, 80
Function field,
  arithmetic automorphic, 364
  smooth model of, 42
Functor,
  exact functor, 212
  fiber functor, 99
  fundamental, 212
  left exact, 212
  Picard, 69, 270
  representability, 72, 75, 209
  representable, 209
  right exact, 212
Fundamental group,
  of $(\mathcal{C}, F)$, 212
  path, 215
  pro-object, 213

Group,
  Barsotti–Tate, 346
  cocenter, 326
  expanding semi-, 165, 228
  formal, 346
  maximal at p, 307
  p-divisible, 346
  Picard, 45
  quasi-split, 218
  special unitary, 305
  split, 218
  symplectic, 224
  types A, B, C, D, 316
  unitary, 224
Group scheme,
  Barsotti–Tate, 346
  cyclic subgroup, 126, 127
  étale cyclic subgroup, 126, 127
  formal, 346
  locally free, 71
  p-divisible, 346

Hasse,
  invariant, 118
  principle, 319
Hecke,
  algebra for $GL(2)$, 177
  character, 176
  module, 7
  nearly ordinary Hecke algebra, 189
  operator, 128, 165
  polynomial, 235
  universal p-nearly ordinary Hecke algebra, 191
Hermitian form,
  anisotropic, 237
  isotropic, 238
  null space, 236
  radical, 236
  signature, 237
  totally isotropic, 238
Hermitian symmetric domain
  bounded realization, 240
  unbounded realization, 240

Igusa,
  curve, 82
  tower, 119, 295, 330
Induction,
  algebrogeometric, 231
  continuous, 333
  regularly induced, 227
  smooth, 226
Involution,
  of the first and the second kind, 304
  positive, 304
Isogeny, 265
  degree of, 265
  étale, 168
  prime-to-Σ, 138
  symmetric, 99
Iwasawa function, 11

Jacobian, 45
Jacquet module, 226

Level structure,
  Γ-structure, 124
  $\Gamma(N)$, $\Gamma_1(\mathfrak{N})$-structure, 102, 103
  $\Gamma_0(\mathfrak{N})$-structure, 127

Modular form,
  central character, 171
  cusp form, 119
  Λ-adic form, 91
  p-adic analytic family, 91
Module,
  of finite corank, 339
  Hecke, 7
  Jacquet, 226
  Tate, 307
Modulus, Moduli, 75
  coarse, 72, 75
  fine, 72
  representability, 72, 75, 209
Morphism,
  epimorphism, 209
  monomorphism, 209
  separable, 216

Nearly ordinary,
  Hecke algebra, 189
  representation, 234
  universal Hecke algebra, 191
Norm,
  reduced, 199
Null space, 236

Object,
  connected, 212
  final, 210
  fundamental pro-object, 213
  initial, 210
  test, 86
Operator,
  diamond, 125, 171, 173
  Hecke, 128, 165

p-adic,
  analytic family of, 91
  ring, 127
Parameter,
  adapted to a differential, 51
Path, 215
Picard,
  functor, 69, 270
  group, 45
  scheme, 69
Plücker coordinate, 256
Polarization, 99, 307
  $\mathfrak{c}$-polarization, 100
  D-linear, 307
  symmetric, 99
Polygon, 235
  Hodge, 235
  Newton, 235
Positive,
  involution, 304
  totally, 100

Radical,
  of Hermitian space, 236
Raynaud extension, 284
Reflex,
  field, 132, 306, 325
Representation,
  admissible, 226
  continuously induced, 333
  nearly p-ordinary, 234
  regularly induced, 227
  schematic, 76, 332
  smooth, 226
  universal, 331
Residue,
  field, 41
  theorem, 48
Riemann–Roch theorem, 49
Rigidity,
  of endomorphisms, 264

Scheme,
  abelian, 98, 264
  ample invertible sheaf on, 276
  Hilbert, 261
  moduli, 75
  ordinary abelian, 101
  Picard, 69
  quasi-projective, 261
  semi-abelian, 116, 118, 265
  stable point of, 273
  strongly projective, 265

strongly quasi-projective, 265
Weil restriction of, 97
Sheaf,
ample invertible, 276
fppf, 346
of cusp forms, 119
Shimura variety,
abelian type, 326
CM point of, 4, 154, 303
geometrically irreducible component of, 141
neutral component of, 142
reflex field, 132, 306, 325
special point of, 325

Tate,
Barsotti–, 346
curve, 78
module, 307
semi-abelian scheme, 116, 118, 265
Test object, 86
Theorem,
Abel's, 56
of Chai–Faltings, 283, 297
of Faltings–Zarhin, 141
residue, 48
Riemann–Roch, 49
vertical control, 9
Trace,
reduced, 199

Uniformizer, 42

Valuation ring, 37
discrete, 37
DVR, 37
uniformizer of, 42

Weierstrass equation, 52
Weight, 5
dominant, 5, 231
double digit, 171
positive, 331
regular, 231
Weil,
pairing, 103, 350
restriction, 97

Zariski topology, 44
Zariski–Riemann space, 39